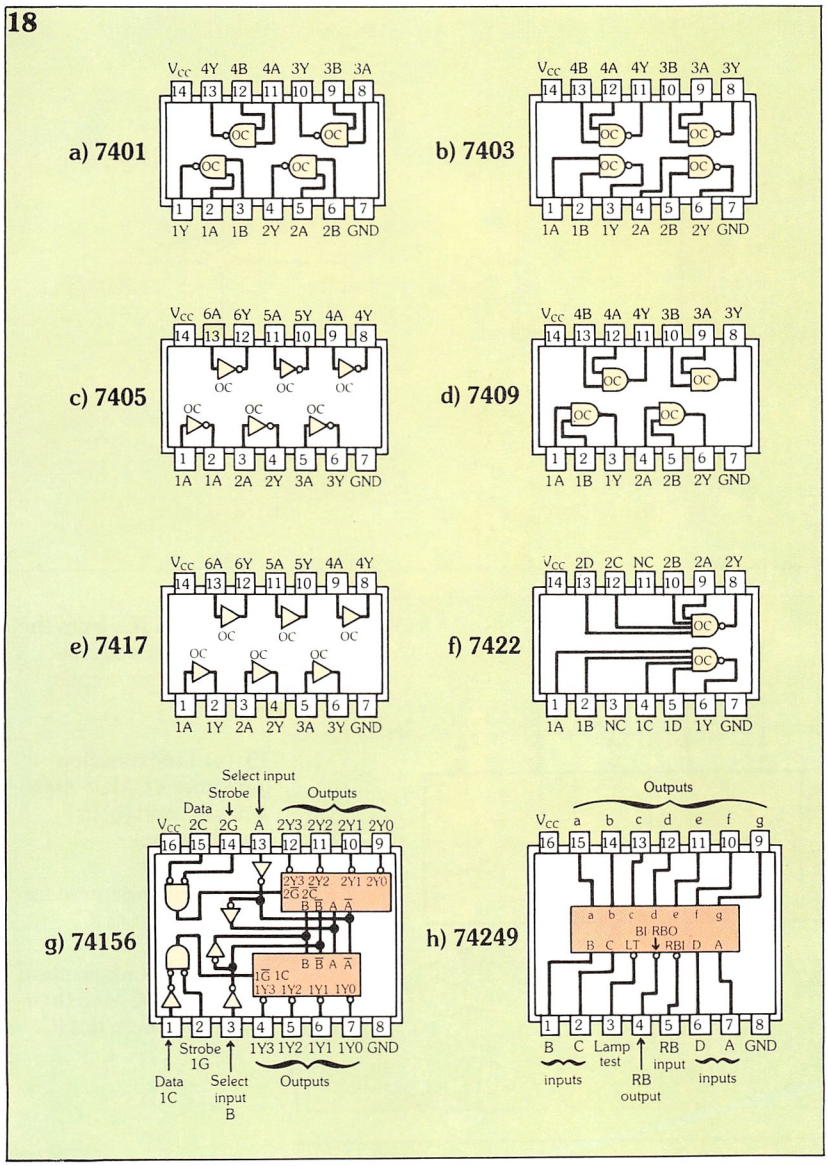
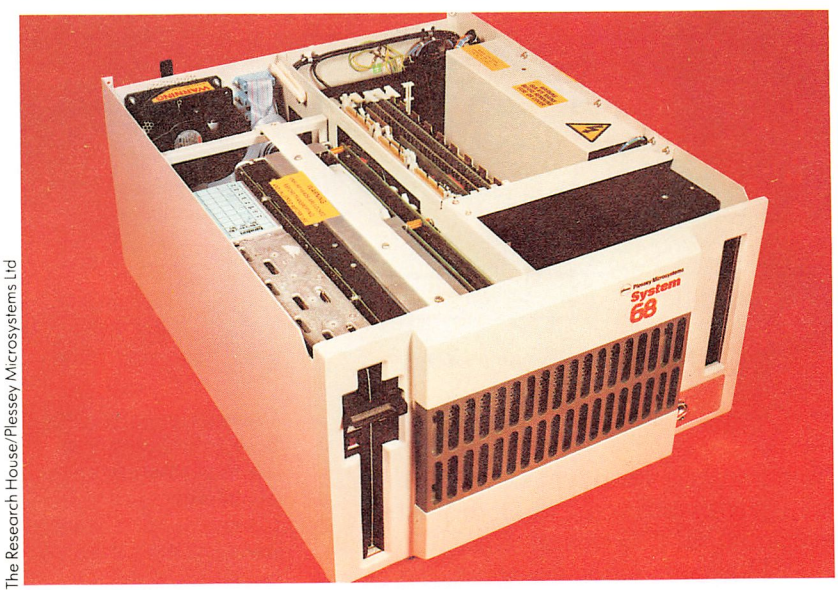


The Research House/Plessey Microsystems Ltd

*(continued from part 17)*

**Examples of TTL open collecter gates**
The following integrated circuits are typical of those from the TTL 74 series with open collector output stages.

**7401 IC NAND gates**
The pin arrangement of this quad, two-input NAND gate IC is shown in *figure 18a*. Each NAND gate in the IC has an open collector output stage.

**7403 IC NAND gates**
The 7403 IC, containing four two-input NAND gates, is functionally similar to the 7401 integrated circuit. However, the pin arrangement (shown in *figure 18b*) is different.

**7405 IC hex inverter**
The pin arrangement of the 7405 IC is shown in *figure 18c*. There are six separate inverters in the IC, all with open collector outputs.

**7409 IC NAND gates**
The pin arrangement of the 7409 IC, quad, AND gate with open collector output stages is shown in *figure 18d*.

**7417 IC hex buffer/driver**
The 7417 IC hex buffer/driver contains six buffer/drivers, all with open collector output stages. Its pin arrangement is shown in *figure 18e*.

**7422 IC NAND gates**
The 7422 IC contains two four-input NAND gates. Its pin arrangement is shown in *figure 18f*.

**74156 IC decoder/demultiplexer**
Open collector output stages are not limited only to ICs containing individual gates. The 74156 IC, for example, is a decoder/demultiplexer with open collector outputs. Its pin arrangement is shown in *figure 18g*.

**74249 IC decoder/driver**
The pin arrangement shown in *figure 18h* is of a 74249 IC. This integrated circuit is a BCD-to-seven-segment decoder/driver with open collector outputs.

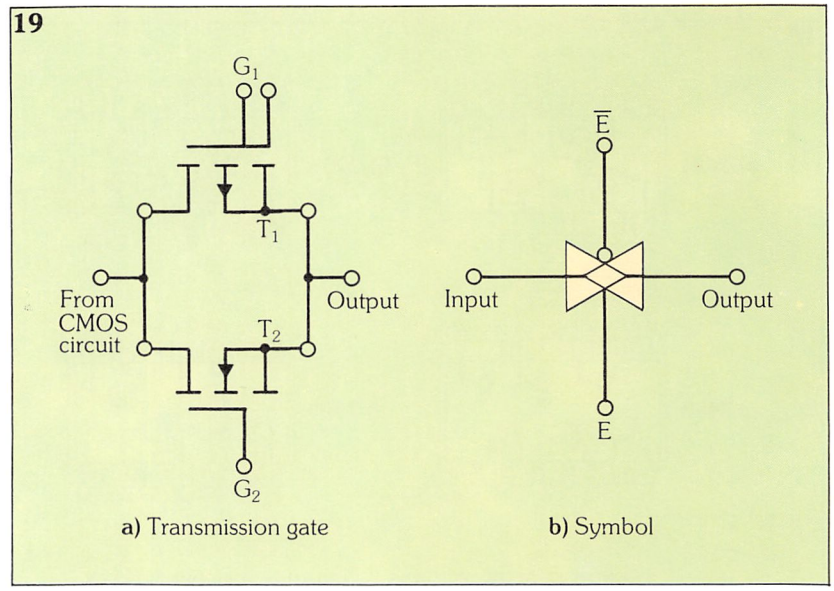
## Three-state CMOS devices

We will now move on to look at some integrated circuits, from the CMOS 400 series, which have three-state outputs. CMOS integrated circuits differ from those of TTL in that they have no identifiable separate output stage, instead, the outputs of the transistors which form the logic circuit are used directly as outputs of the IC. Therefore, to make a three-state CMOS device a separate stage, known as a **transmission gate** and shown in *figure 19a*, has to be added to the circuit. The transmission gate is made from one p-channel and one n-channel MOSFET (transistors $T_1$ and $T_2$ in *figure 19a*) connected as shown. The logic symbol for a transmission gate is shown in *figure 19b*.

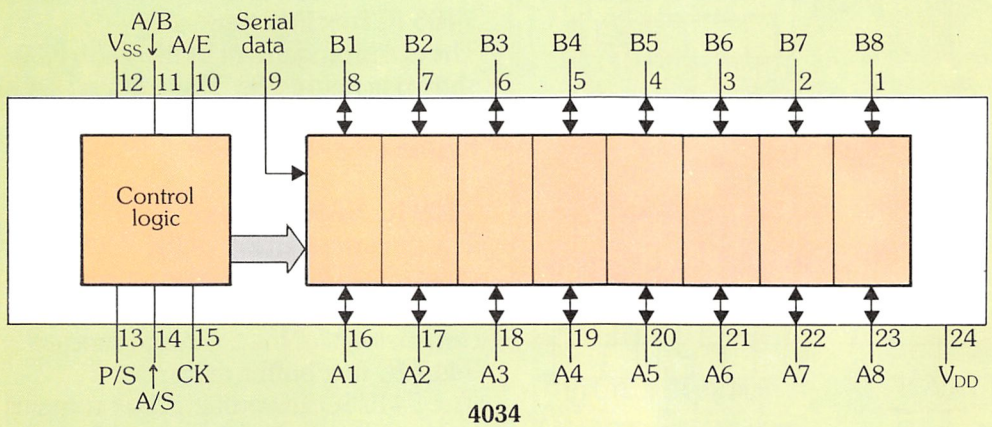
When the gate $G_1$, of p-channel MOSFET $T_1$, is at logic 0, and the gate $G_2$, of n-channel MOSFET $T_2$, is at logic 1, both transistors conduct and the transmission gate connects the input (from the preceding CMOS circuit) to the IC output.

However, when the gate $G_1$, is at

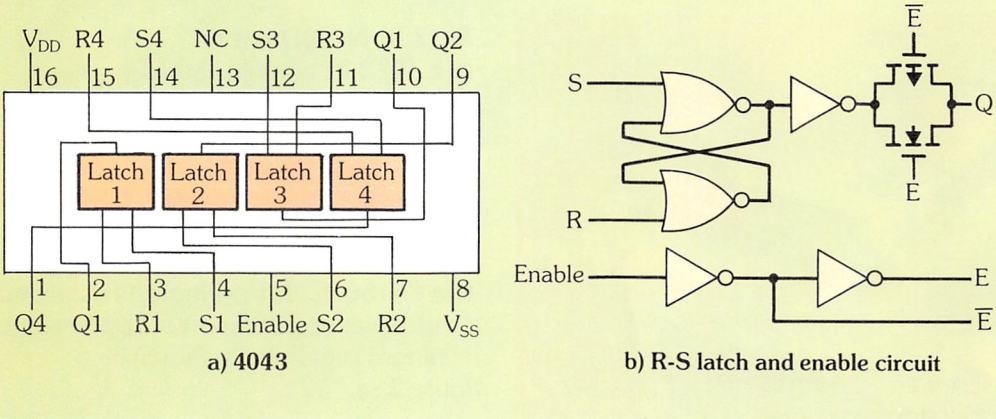
19




20




21




**18. Typical ICs from the TTL 74 series** all with open collector output stages.

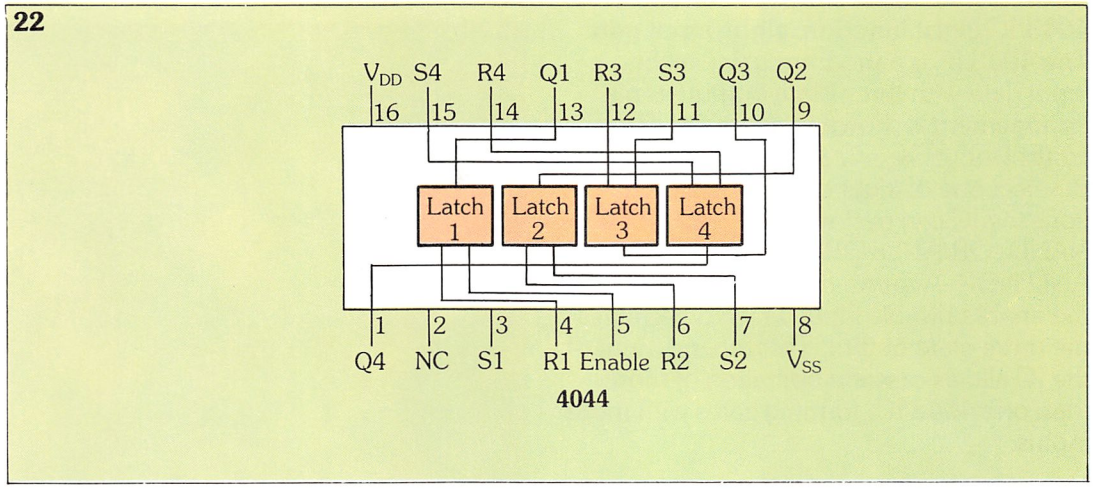
**19. (a) Transmission gate** from a CMOS 400 series IC; and **(b)** its circuit symbol.

**20. Pin arrangement** for the 4034 CMOS IC.

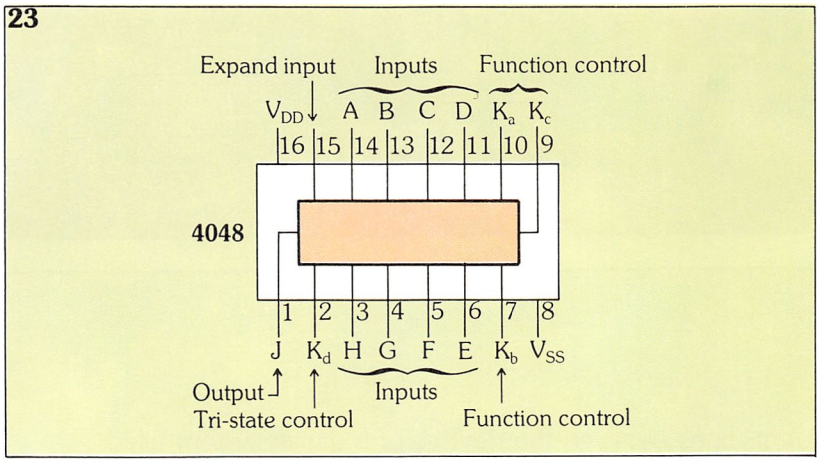
**21. (a) Pin configuration** of the 4043 IC; and **(b)** a single latch from this IC.

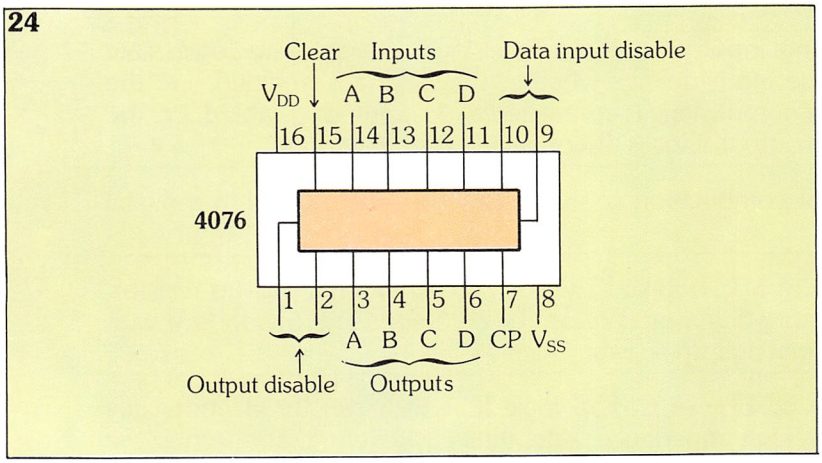
**Previous page:** a UNIX based microcomputer which can handle up to 16 different users. Communication between users and the CPU is via busses.

**22. Pin arrangement** for the 4044 IC.




**23. Pin arrangement** for the 4048 IC.




**24. Pin configuration** for the 4076 IC quad, three-state D-type flip-flop.

logic 1, and the gate $G_2$ is at logic 0, neither MOSFET conducts, and the output is disconnected from the input, and the transmission gate presents a high impedance output state.

The complementary signals required at the gates of the MOSFETs in the transmission gate are used as the enable/disable (E and $\overline{E}$) inputs to the three-state CMOS device. These inputs are generally joined with an inverter allowing use of a single controlling signal.

The following integrated circuits are examples of CMOS three-state ICs from the 4000 series.

### 4034 IC bidirectional bus register

The 4034 CMOS IC is an 8-bit shift register which can be used to transfer data bidirectionally between two busses; its pin arrangement is shown in *figure 20*. The device can convert serial data to parallel data, or vice versa, with data transfer after conversion to either of the connected busses.

One bus is connected to input/output terminals A1 to A8; the second bus is connected to terminals B1 to B8. All bus input/output terminals have three-state outputs.

### 4043 IC NOR gate R-S latch

*Figure 21a* shows the pin arrangement of the 4043 IC quad, three-state NOR R-S latch. The IC features a common enable/disable line which controls the three-state output of all four latches simultaneously. A single latch of the IC is shown diagrammatically in *figure 21b*, along with the circuit producing the E and $\overline{E}$ enable connections to the gates of the two MOSFETs in each transmission gate.

### 4044 IC NAND gate R-S latches

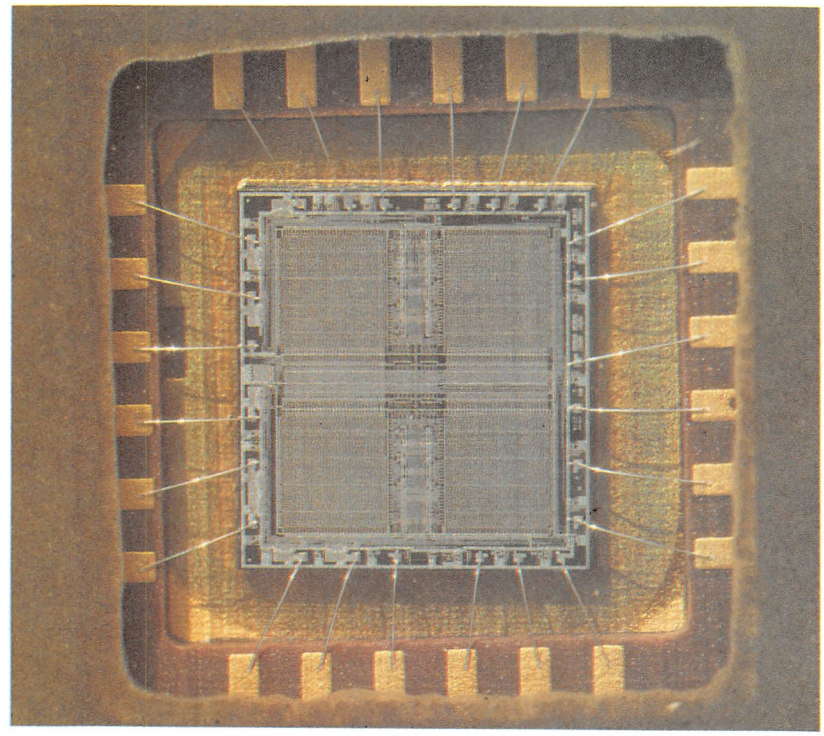
The 4044 IC (pin arrangement in *figure 22*) is functionally similar to the 4043 IC, but the four latches are made from NAND gates.

### 4048 IC eight function, eight-input gate

The 4048 IC is a programmable eight-input gate with three-state output; its pin arrangement is shown in *figure 23*. Three control inputs $K_a$, $K_b$ and $K_c$ can be used to select one of eight different logic functions the IC can perform: OR, NOR, AND, NAND, OR/AND, OR/NAND, AND/OR, AND/NOR. A fourth control input $K_d$, is the enable/disable input which determines the three-state output. The expand input of the IC allows cascade operation of more than one 4048 IC, forming gates with more inputs.

### 4076 IC D-type flip-flops

The pin arrangement of the 4076 IC quad, three-state D-type flip-flop is shown in *figure 24*. A common clock operates all four flip-flops synchronously and they are positive-edge triggered. A clear input resets the flip-flops when a logic 1 signal is applied.



ZEFA

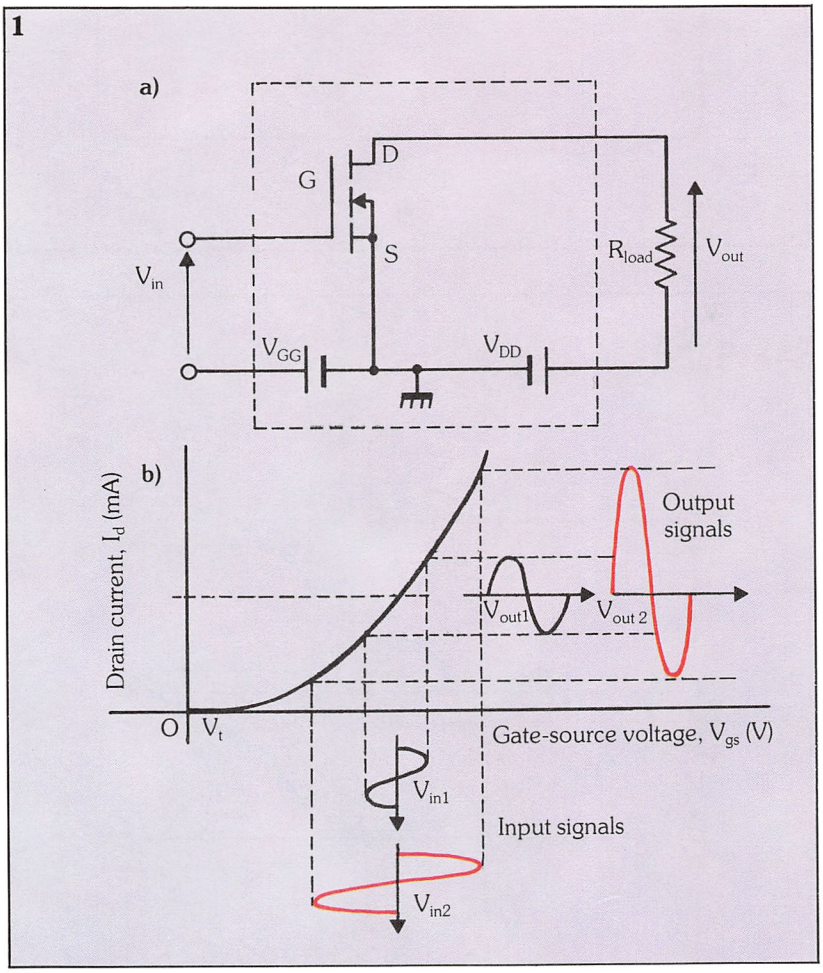
## Glossary

| | |
|---|---|
| **bus** | a group of connections between the devices of a digital system used for transfer of data |
| **enable/disable input** | the controlling input of a three-state logic device. The logic state applied to the input defines whether the device is enabled, i.e. the output of the output stage is connected to its input, or disabled, i.e. the output of the output stage is disconnected from the input |
| **illegal** | logic state, or combination of states, which is not allowed in a digital system |
| **open collector output state** | output stage of a TTL logic IC which has no internal pull-up resistor. Outputs of more than one gate can be combined and tied to + V with a single external pull-up resistor |
| **three-state output stage** | output stage of TTL or CMOS logic IC which can be electronically switched to a high impedance state, thus effectively disconnecting the IC's internal circuit from the output terminals |
| **transmission gate** | part of a CMOS logic IC which can be electronically switched between a high and low impedance. Used in three-state CMOS devices |

17

SOLID STATE ELECTRONICS

# Analogue MOSFET circuits

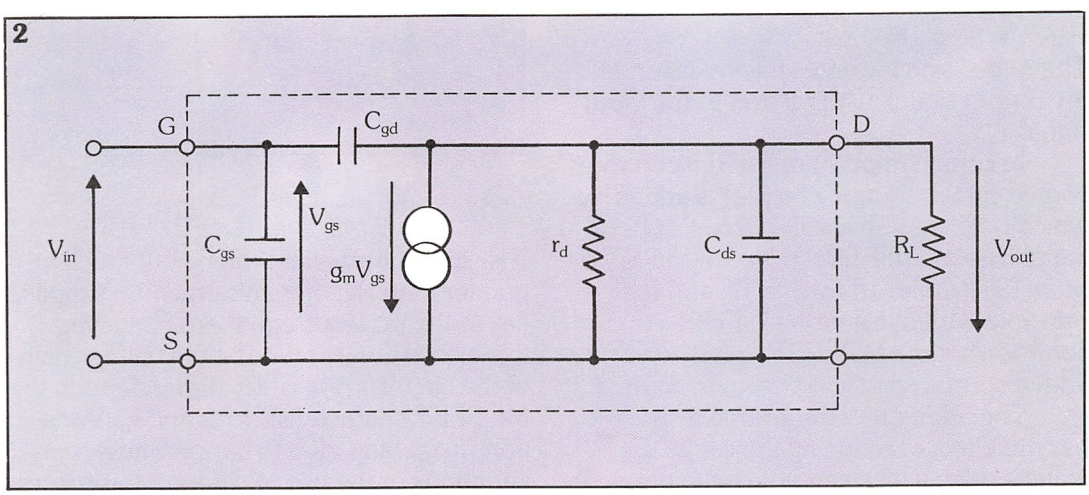



## Using equivalent circuits

As we have seen in *Solid State Electronics 14*, equivalent circuits serve as useful tools in the analysis of complex semiconductor circuits, however, they can only be used if the complex circuits are *linear* in operation. Although both JFETs and MOSFETs are inherently non-linear (as their transfer characteristic curves illustrate) they can be considered as approximately linear devices when only small signals are amplified. Thus, we are justified in the use of small signal equivalent FET circuits.

The effect of FET non-linearities is shown in *figure 1a* – here the circuit of an n-channel enhancement-type MOSFET has an applied small-signal input signal, $V_{in}$. The gate-source bias voltage is provided by a battery of voltage $V_{GG}$, and the resulting output voltage, $V_{out}$, is an amplified version of the small-signal input voltage. A typical transfer characteristic curve of an n-channel enhancement-type MOSFET is shown in *figure 1b* and two values of input voltages have been added to it. The first applied input voltage, $V_{in1}$ (a small input voltage), produces the output, $V_{out1}$, which we can see is amplified and not noticeably distorted. However, the larger

**1. (a) The circuit of an n-channel enhancement-type MOSFET** with an applied small signal input signal; **(b)** typical transfer characteristics.




**2. Equivalent circuit** for central block of *figue 1a* (as shown by the dotted lines).

**Above left:** 2716 PROM capable of storing 2048 8-bit words.

second applied input voltage, $V_{in2}$, produces the output $V_{out2}$, which *is* noticeably distorted.

**MOSFET small-signal circuits**

This approximation to a linear circuit allows the part of the circuit shown in the central block of *figure 1a* to be replaced by the small-signal equivalent circuit of a MOSFET (*figure 2*). Capacitive effects within the body of the MOSFET (known as **parasitic capacitances**) have been shown as three capacitors: between gate and source ($C_{gs}$), between gate and drain ($C_{gd}$), and between drain and source ($C_{ds}$). Although the exact values of these capacitors depend on the type of device being used, capacitances of between 1 to 5 pF are common. As the reactance of a capacitor falls with increasing frequency, we can see from the equivalent circuit of *figure 2* that the output voltage of the circuit will fall at high frequency. For low frequency applications however, these parasitic capacitors can generally be ignored.

The output voltage of this circuit at low frequencies is:

$$V_{out} = -g_m V_{gs} (R_L // r_d)$$

where $R_L // r_d$ means the overall resistance of the load resistor $R_L$, in parallel with the dynamic drain resistance, $r_d$. The voltage gain $A_v$, of the circuit is therefore:

$$A_v = \frac{V_{out}}{V_{in}}$$

$$= -g_m (R_L // r_d)$$

because $V_{gs} = V_{in}$.

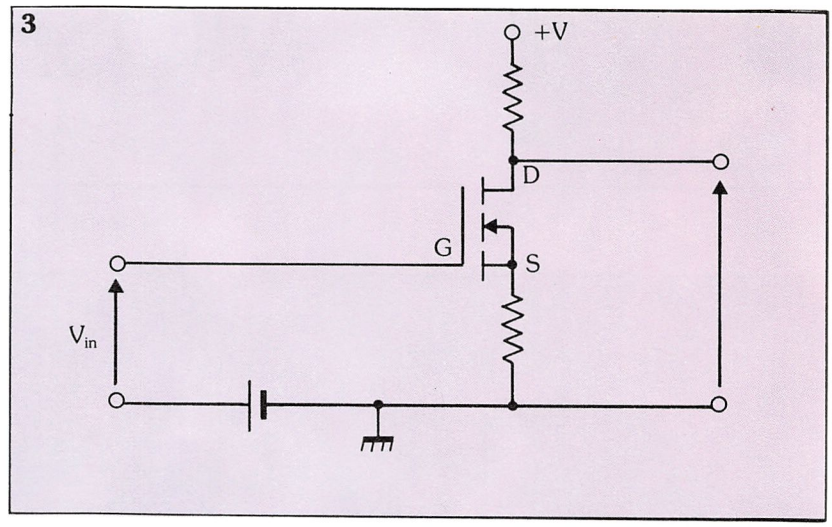
If the dynamic drain resistance is much greater than the load resistor the voltage gain is simply:

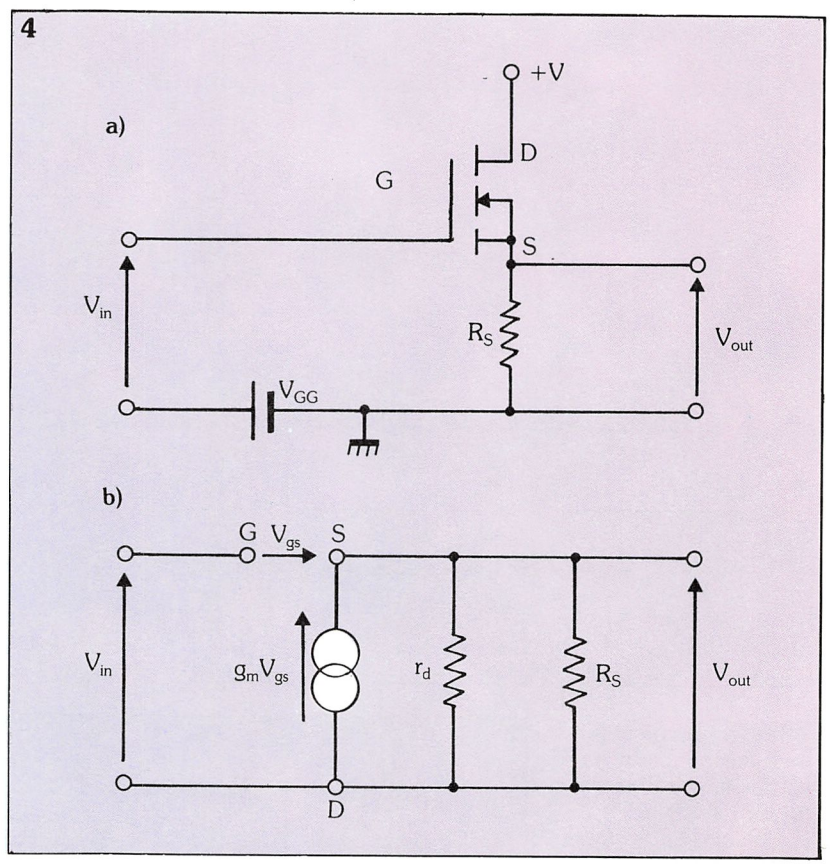
$$A_v = -g_m R_L$$

The minus sign indicates the inversion of the output signal, with respect to the input signal.

In certain applications of the common source circuit, a source resistor is added which helps to stabilise the circuit's operating point – we saw how this works in *Solid State Electronics 14* for JFETs, and the principle is identical for MOSFETs. A common source MOSFET circuit, with added source resistor is shown in *figure 3*.

The addition of the source resistor has the effect of reducing the circuit's voltage gain, which can now be expressed:




$$A_v = -\frac{g_m}{1 + g_m R_S} (R_L // r_d)$$

If $R_L$ is much smaller than $r_d$, but much greater than $R_S$, then this equation simplifies to the previous equation. However, this expression shows that the voltage gain of the circuit is not totally dependent on the MOSFET characteristics $g_m$ and $r_d$. Variations in $g_m$, say, due to temperature variations or the use of different transistors,

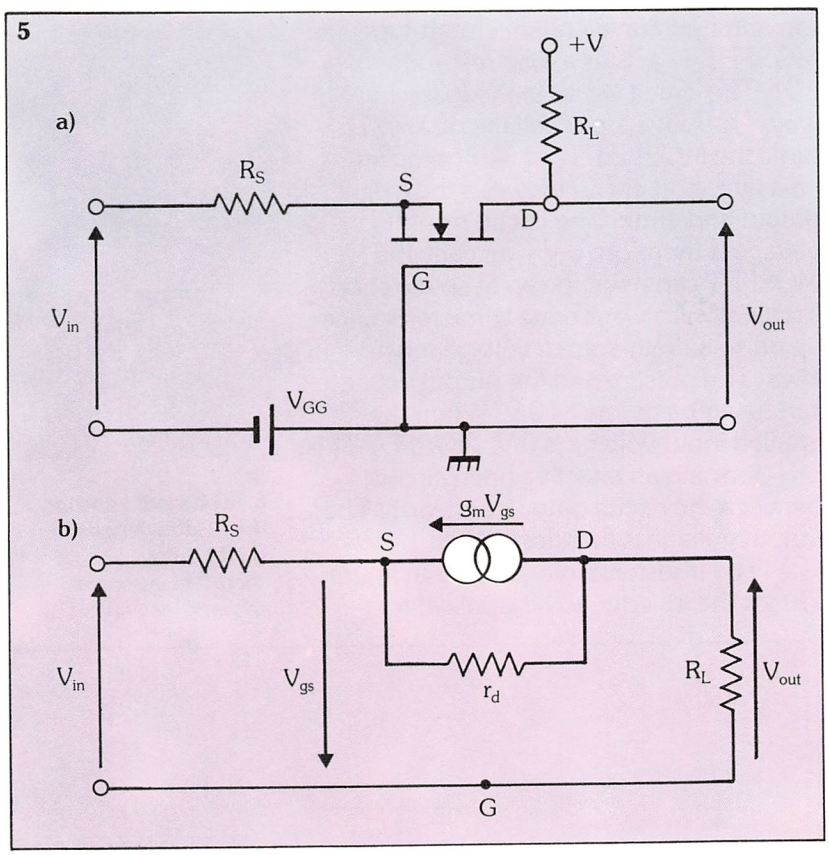
**3. A common source MOSFET circuit**, with added resistor.

**4. (a) An n-channel enhancement-type MOSFET** in a common drain configuration amplifier; **(b)** small-signal equivalent circuit.

will affect both numerator and denominator in a similar way, so, if $r_d$ is much greater than the load resistor the voltage gain can be approximated to:

$$A_v \approx -\frac{R_L}{R_S}$$

The approximate voltage gain of the circuit is therefore dependent not on the parameters of the MOSFET, but on the fixed ratio of two resistors.




**5. (a) Common gate MOSFET circuit; (b)** equivalent circuit.

**Right:** a two stage GaAs FET amplifier. It has isolators at input and output, and is protected against over-voltage and polarity reversal.

### Common drain MOSFET circuits

The circuit of an n-channel enhancement-type MOSFET in a common drain (or source follower) configuration amplifier is shown in *figure 4a*. If we restrict this circuit's operation to low frequency, the parasitic capacitances within the MOSFET can be ignored and so a small-signal equivalent circuit looks like that in *figure 4b*, where the source resistor $R_S$, is effectively in parallel with the dynamic drain resistance $r_d$.

Voltage gain of the circuit is given by:

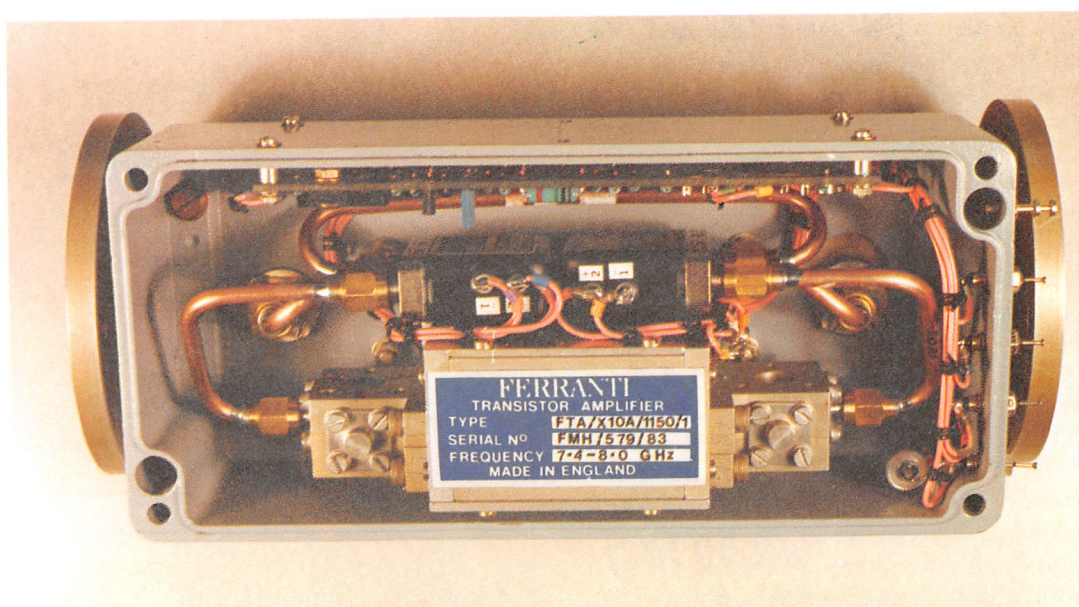
$$A_v = \frac{g_m (R_S // r_d)}{1 + g_m (R_S // r_d)}$$

When the dynamic drain resistance is much larger than the source resistor, the expression simplifies to:

$$A_v \approx \frac{g_m R_S}{1 + g_m R_S}$$

which, for large values of $g_m R_S$, becomes a maximum of unity. Generally, voltage gains of just below this are possible – for example if a MOSFET has a transconductance of 1000 $\mu$S and a source resistor of 10 k$\Omega$, the voltage gain of a common drain configuration amplifier will be approximately 0.9.

The common drain MOSFET amplifier has a low output impedance (typically




The Research House/Ferranti Electronics Ltd

about 500 Ω to 2 kΩ) and a very high input impedance (typically over 1 MΩ). This, coupled with its near unity gain and the fact that output and input signals are in phase, make the circuit useful as a buffer stage between high and low impedance parts of a circuit.

### Common gate MOSFET circuits

A common gate MOSFET circuit is shown in *figure 5a*, and an equivalent circuit is shown in *figure 5b*. Input impedance of the circuit is very low, while its output impedance is high. No phase reversal takes place between input and output, and the voltage gain is approximately:

$$A_v \approx g_m R_L$$

### Self bias

In the circuit shown in *figure 6a*, the gate of an n-channel enhancement-type MOSFET has been biased by connecting a resistor, $R_f$, between the transistor's drain and gate. Since virtually zero gate current is required to operate the transistor, the voltage drop across the resistor (from Ohm's law) is zero, and so the gate-source voltage is equal to the drain-source voltage.

*Figure 6b* shows output characteristic curves of a MOSFET, along with a static load line. Also on the graph is a curve corresponding to $V_{GS} = V_{DS}$. The point at which this curve intersects the load line is the operating point of the circuit, point P. We can effect changes in the operating point by altering the value of the load resistor $R_L$, hence changing the slope of the load line. However, changing the value of resistor $R_f$ has no effect on the operating point.

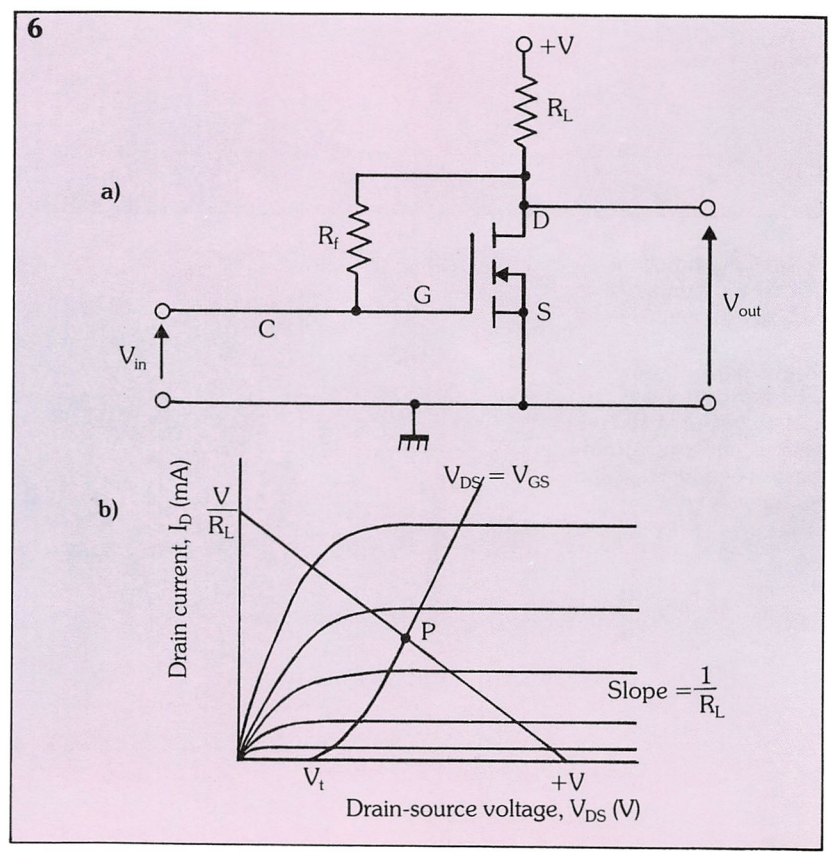
The presence of the resistor between drain and gate helps to stabilise the operating point of the circuit in the following way. If, say, the drain current falls, the drain-source voltage becomes more positive; because the gate-source voltage always equals the drain-source voltage, it also becomes more positive, turning the MOSFET on further; this, in turn, increases the drain current, in a self-regulating manner. The concept of feeding part of the output of a circuit back to the input is known as **negative feedback** and resistor $R_f$ of this circuit is known as a **feedback resistor**.

## Large-signal MOSFET amplifiers

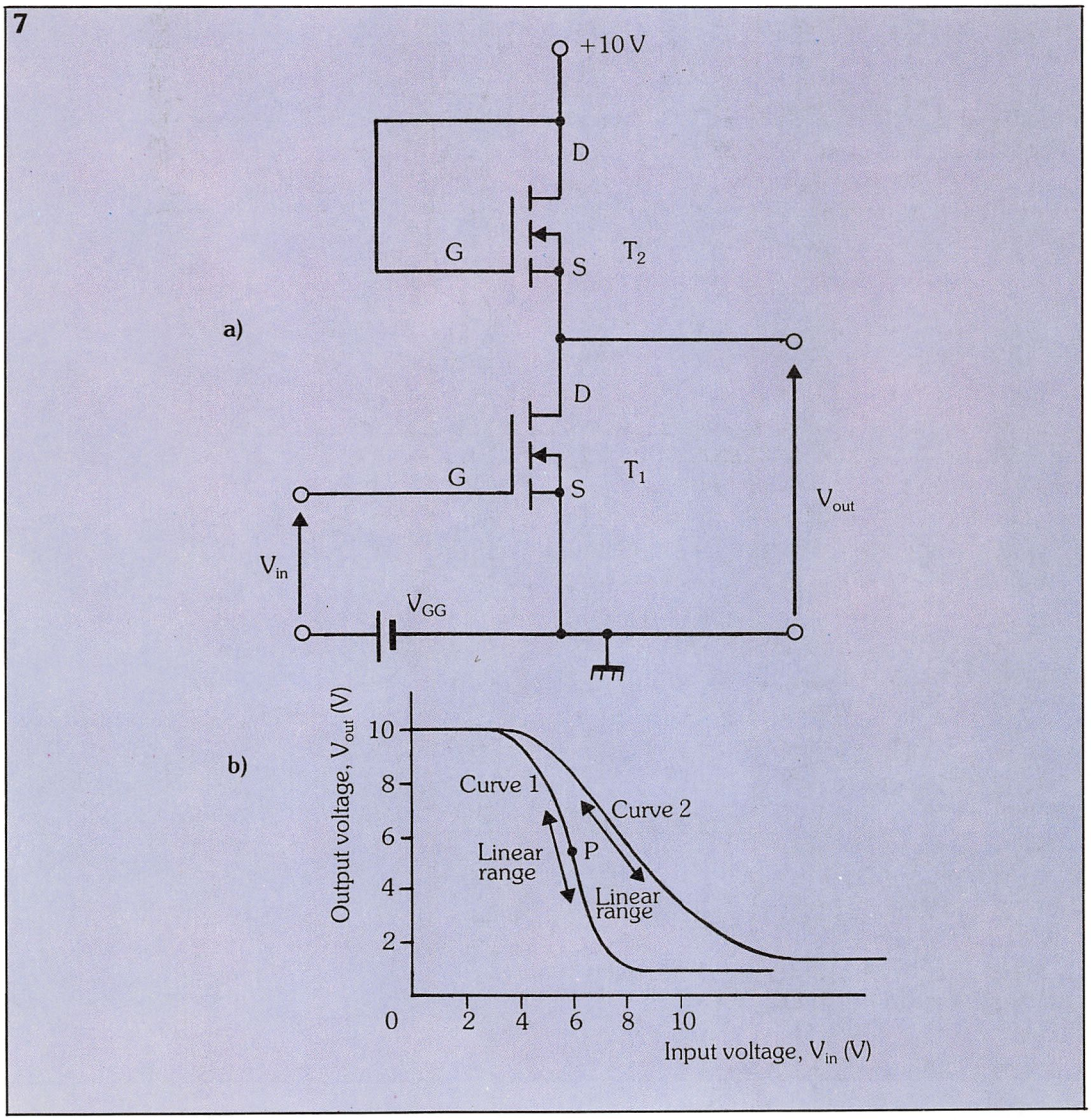
In *Digital Electronics 3* we saw how an inverter (or NOT gate) can be formed from two MOSFETs in series. *Figure 7a* shows a circuit in which n-channel enhancement-type MOSFET $T_1$ is the controlling transistor, and n-channel enhancement-type MOSFET $T_2$ acts as a load resistor.

The circuit works in the following way. A positive input voltage of, say, 10 V turns the MOSFET $T_1$ on – forming an effective short circuit between the circuit output and earth. The circuit's output voltage is therefore 0 V – actually the MOSFET can never be an absolute short circuit, it will always have *some* resistance – and so a drain-source voltage must always be present and the output voltage can never be as low as 0 V. When the applied input voltage is 0 V, MOSFET $T_1$ is off – forming an effective open circuit between the circuit output and earth. The output voltage is therefore about 10 V.

The transfer curve (curve 1 in *figure 7b*) for the inverter in this application




**6. (a) Circuit showing the feedback resistor, $R_f$; (b)** output characteristic curves showing load line.




**7. (a) Two MOSFETs in series** forming an inverter (NOT gate); **(b)** transfer curve showing how output voltage varies with input voltage.

shows how the output voltage, $V_{out}$, varies with the input voltage $V_{in}$, over the whole range from 0 to 10 V. The transition in output voltage between 10 V and approximately 0 V, for a corresponding change in input voltage of about 3 V to 6 V, occurs over a variable section of the transfer curve. If we fix the circuit's operating point as, say, point P on curve 1 (a part of the curve which is almost linear) then the circuit can be used as an inverting amplifier for small-signals.

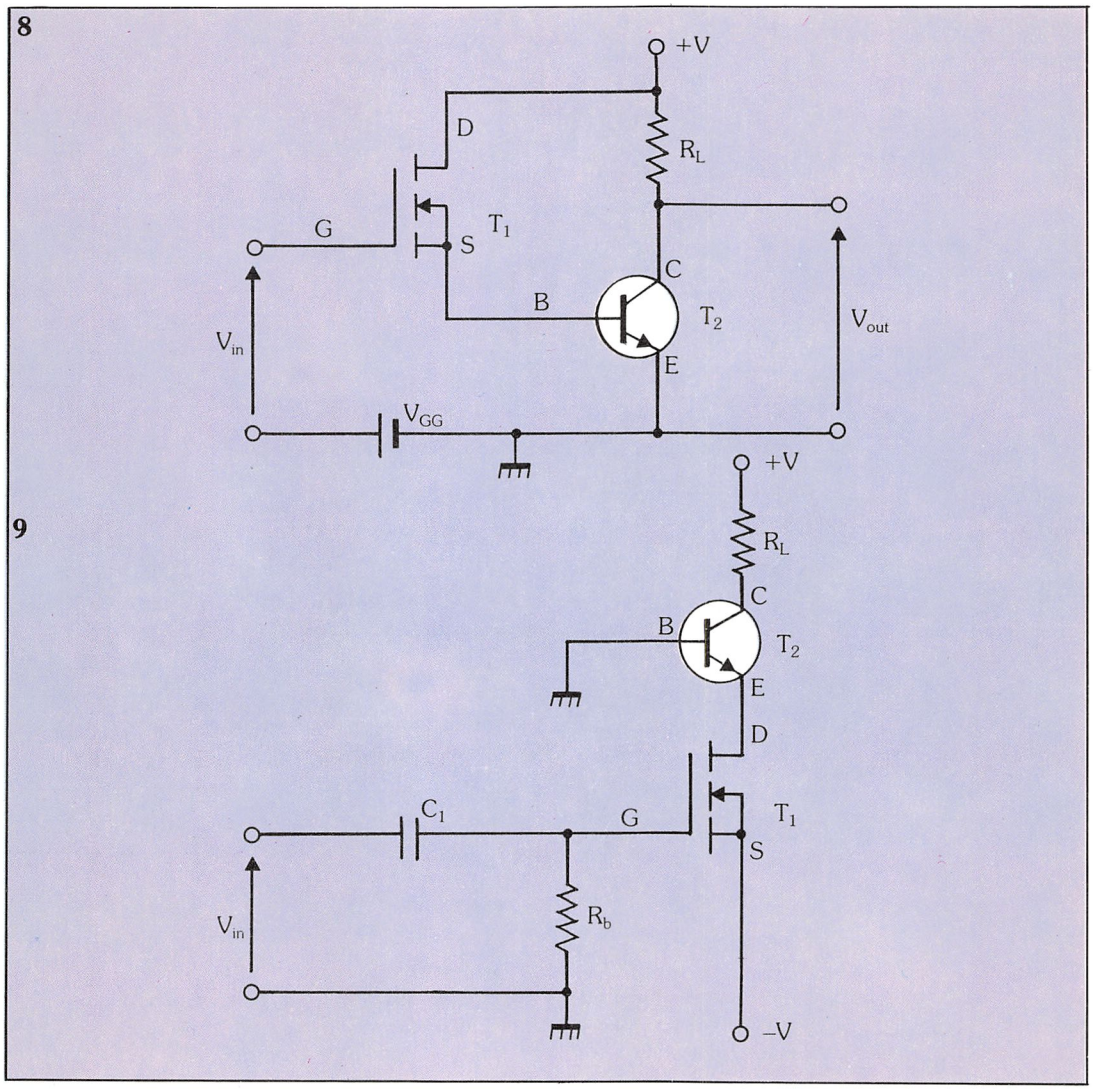
We can increase the length of this linear region of the transfer curve (and therefore increase the size of the signals which can be amplified) by altering the gain of the circuit. Curve 2, for example, shows the transfer curve of an inverter circuit where the linear region has clearly been extended. The gain of the circuit can be given by the expression:

$$A_v = -\sqrt{\frac{(W/L)_{T2}}{(W/L)_{T1}}}$$

where W is the width of each MOSFET channel and L is each MOSFET channel's length. So by making transistors with required channel dimensions we can define the gain of the circuit, and in so doing, define the operating point and the maximum size of signals (now large-signals) which may be amplified.

### MOSFETS and bipolar transistors

Extremely high gains can be obtained by combining a MOSFET with a bipolar transistor amplifier. For example, the circuit of *figure 8* shows an n-channel en-




**8. An extremely large gain** may be obtained from this circuit where a MOSFETand a bipolar transistor are combined in a Darlington pair configuration.

**9. An inverting amplifier** made by combining a MOSFET and a bipolar transistor.

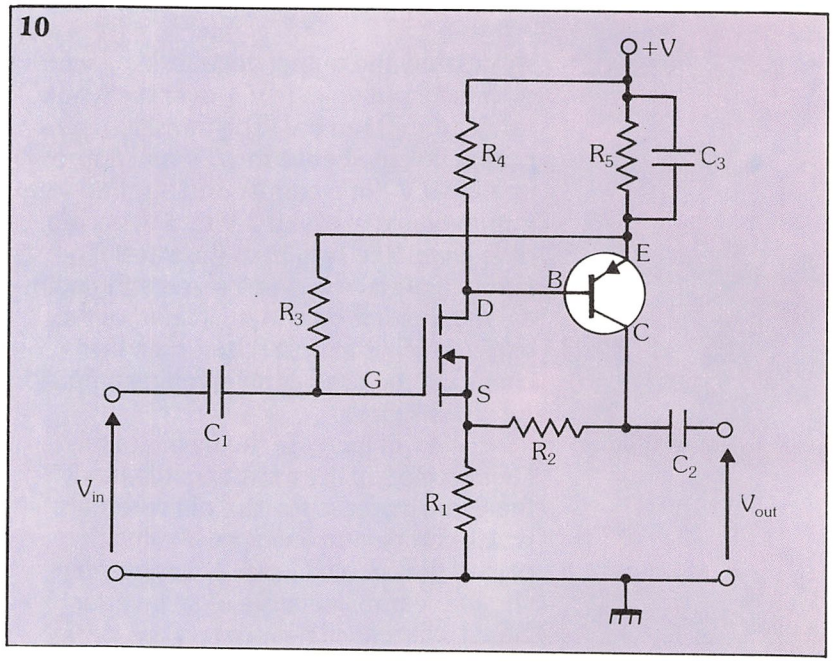
**10. Negative feedback** in a MOSFET, bipolar transistor configuration.

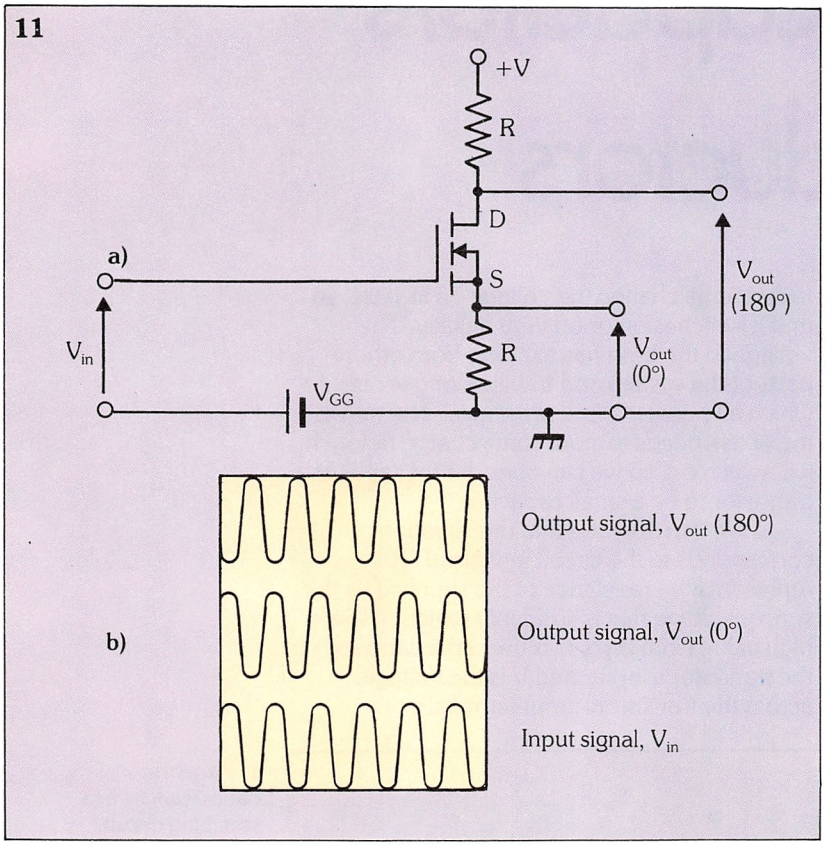
**11. (a) Phase splitter circuit**; **(b)** output signal waveforms from such a circuit.

hancement-type MOSFET and an n-p-n bipolar transistor in a Darlington pair configuration. The overall gain of the two devices is equal to the product of the gains of the individual transistors. The desirable properties of both types of transistors are combined by such a circuit. In this way an extremely high input impedance amplifier, with a large gain, may be formed.

The circuit arrangement shown in *figure 9* is another way of combining MOSFETs with bipolar transistors. Here, an inverting amplifier, similar to that of *figure 7*, has been made with a bipolar transistor as transistor $T_2$.

*Figure 10* shows an arrangement with a MOSFET and a bipolar transistor which features negative feedback. When the individual transconductances of the two transistors are large, the overall circuit gain






is independent of the parameters of the devices, and can be approximated by:

$$A_v = \frac{R_1 + R_2}{R_1}$$

The circuit gain is thus controlled by the values of passive components.

**Phase splitters**

A MOSFET may be used as the basis of a circuit which provides two output signals in antiphase (i.e. with a phase difference of 180°) from a single input signal. *Figure 11a* illustrates a basic configuration. The amplitude of the MOSFET's drain current is the same as the amplitude of the source current. The two currents do, however, differ in phase by 180°. Bipolar transistors are frequently used in such an arrangement although the size of a bipolar transistor's collector current is not exactly the same as its emitter current: these two voltages may be made equal by making the two resistors unequal.

The graph in *figure 11b* shows the output signal waveforms which may be obtained from such a **phase splitter** circuit.

# Glossary

| | |
|---|---|
| **feedback resistor** | a resistor used to feed part of a circuit's output back to its input, normally to aid the circuit's operation in some way. Often the value of the feedback resistor determines the circuit's gain |
| **large-signals** | a signal waveform of several volts amplitude |
| **negative feedback** | circuit principle in which part of the output signal of a circuit is fed back to the circuit input – the overall effect is to reduce gain and increase stability. In many cases, negative feedback may be used to accurately define a circuit's gain |
| **parasitic capacitances** | capacitive effects occuring within an electronic device which can affect the device's operation at high frequency. Parasitic capacitances with a MOSFET decrease a transistor's effective gain at high frequencies |
| **phase splitter** | a circuit used to form two antiphase output signals from a single input signal. May be formed using a single MOSFET |
| **small-signals** | signals of small amplitude. An essentially non-linear MOSFET amplifier may be used to amplify small-signals because the transistor is not operated out of the limited linear range |

BASIC THEORY REFRESHER

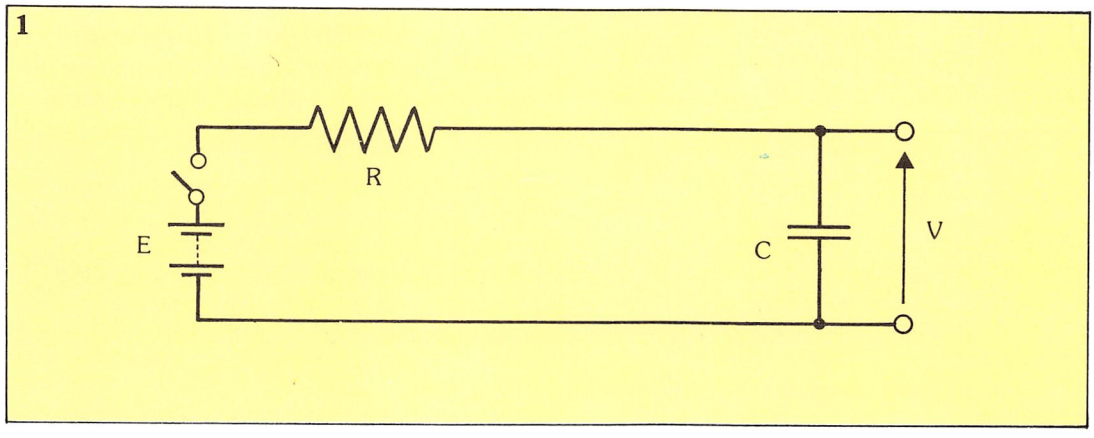
ELECTRICAL TECHNOLOGY

# Charge in capacitors and inductors

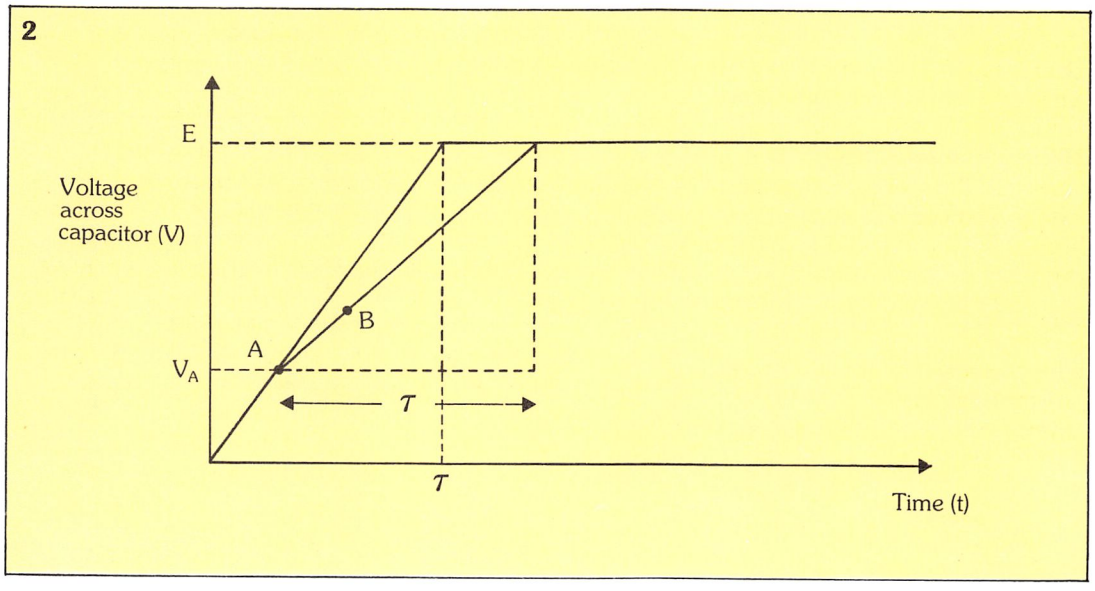
Up to now, we have considered capacitors to be discrete components primarily comprising two metal plates separated by a dielectric. However, it is important to realise that a capacitance can exist wherever we have a group of conductors separated by air or another insulating material. These **parasitic** capacitances are usually small enough to be neglected, but they may occasionally become significantly large and affect the operation of electronic circuits. In a similar way, inductors exist everywhere that there are wires carrying currents.

Let's take the case of a transistor that is being used as a switch in a computer. We have to suddenly change the voltage on its base, so that it switches on or off very quickly. The voltage to the base has to come from other parts of the system and travels along wires which all possess some resistance. The transistor's base needs to store some charge before it will switch on, so we can consider the input of a transistor to be a small capacitor.

We can now see that this situation corresponds to the circuit in *figure 1*. R represents the resistance of the supply; E is the supply voltage that is suddenly switched to a high value from zero; C is the capacitance of the transistor's input; and V is the voltage across the transistors input terminals.

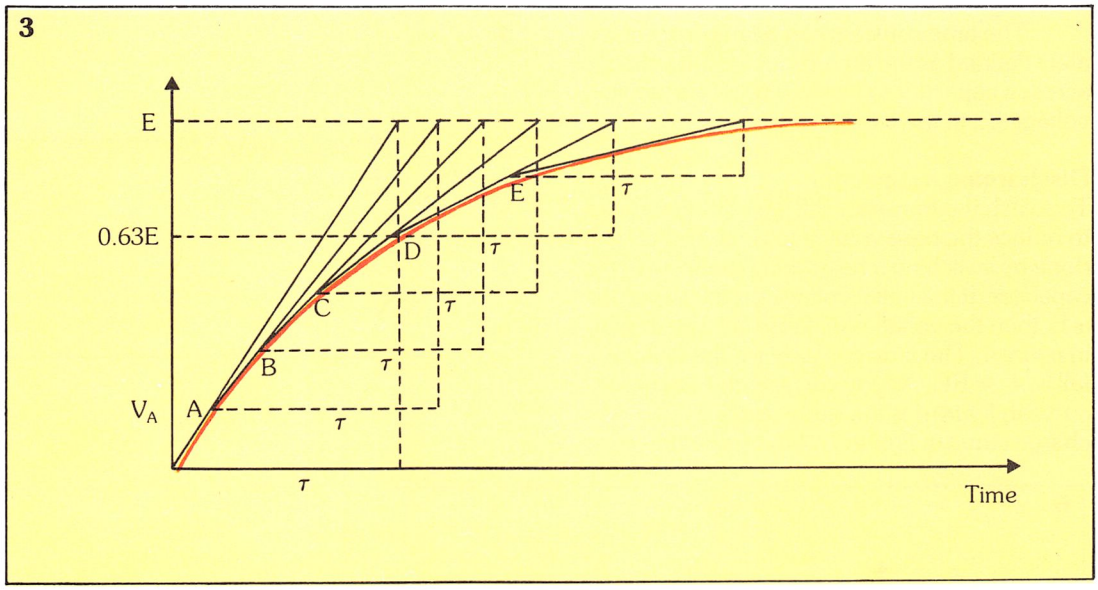



**1. Parasitic capacitances** in a switching circuit.

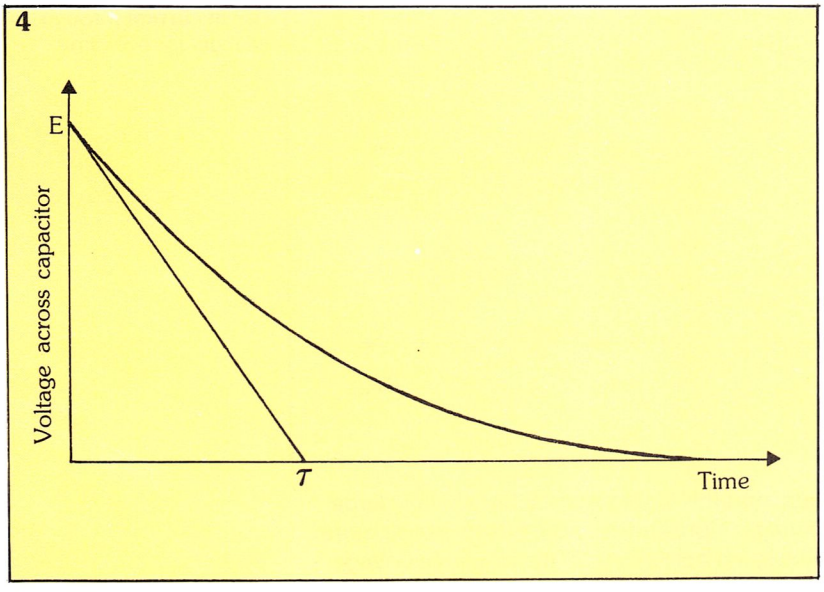



**2. Graph illustrating rising voltage** across the capacitor with time, for the circuit shown in *figure 1*.

**3. Charging a capacitor** – the voltage rises and approximates its final value after time 5 $\tau$ .




**4. Discharging the capacitor** – the shape of the curve is the mirror image of the charging curve of *figure 3*.




## Charging a capacitor

We ideally want V to change from 0 to E volts almost instantaneously. This type of problem is typical of many we may meet in electronic circuits. At the moment the switch is closed, the voltage V will be zero and a current, I, will flow through R. The magnitude of this current is given by:

$$I = \frac{E}{R}$$

If this current lasts for time t, the charge transferred onto the capacitor will be I x t, since current is the rate of flow of charge.

The voltage, V, on the capacitor is given by:

$$V = \frac{Q}{C}$$

$$= \frac{It}{C}$$

$$= \frac{Et}{RC}$$

When the capacitor is fully charged its voltage cannot be bigger than the supply, E. So after a time $\tau$ , the capacitor is fully charged to E, and:

$$E = \frac{E\tau}{RC}$$

$\tau$ is known as the time constant of the circuit and is equal to RC. This is indicated by the red line in *figure 2*. However, the voltage *cannot* actually rise along this line. When the voltage has risen to say, point A (where the voltage across the capacitor is $V_A$), the voltage driving the charge through the resistor will be $E-V_A$. This means that the charge will not be flowing as fast, and the voltage rise across the capacitor will be correspondingly slower.

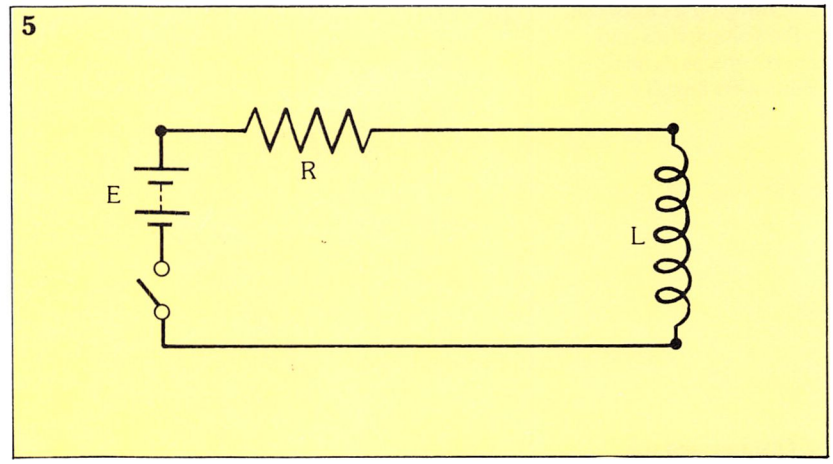
At this point we can see that we are now trying to charge up the capacitor from the voltage $E-V_A$ to the final value E – which it will reach after time $\tau$ . This is shown in green in *figure 2*. Again, we can see that this will only hold for a short period of time, say to B. The effective supply voltage will then be $E-V_B$. If this succession of events is continued, and the steps A, B, C etc. are taken very close together, we will get a smooth curve like the one in *figure 3*. This shows the way that voltage across the capacitor rises towards its final value. We can see that it will never reach the value E, but it will be sufficiently close to it for all practical purposes after a time of about 5 $\tau$ .

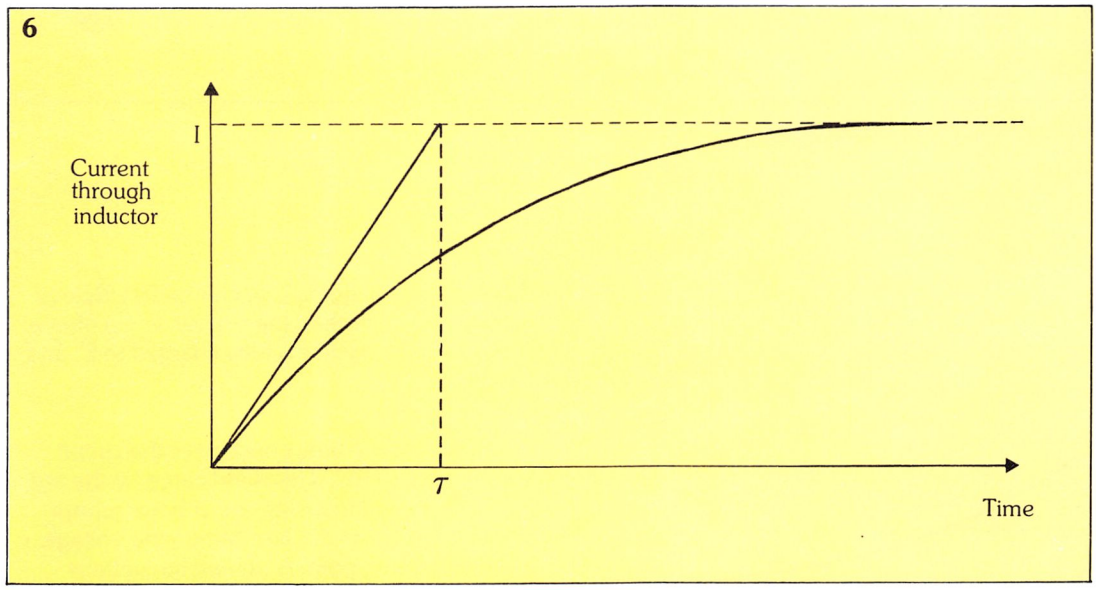
Returning to our problem of the transistor we see that the smaller the time constant, the faster the rise in its base voltage. We should therefore keep R and C as small as possible.

The time constant can be more accurately defined as the time taken for the voltage across a capacitor to rise to 0.63 of the applied voltage (E, in this case).

**Discharging a capacitor**

To switch the transistor off, we obviously need to reduce the base voltage to zero. This can be done by switching a resistor in parallel with the capacitor. If the initial voltage on the capacitor is E, then the voltage will fall to zero as shown in *figure 4*. The time constant for the voltage fall is, $\tau = RC$. As you can see, the discharge curve in *figure 4* is the same shape as the charge curve in *figure 3* – but upside down.




**5. Switching circuit** with an inductor in series.

**6. Curve showing the rise in current** through an inductor over time.

We now know that the time constant is a very important parameter used to indicate the speed that a capacitor charges or discharges through a resistor.

**Build-up of flux in an inductor**

*Figure 5* shows a circuit with a power source, a switch, a resistor and an inductor, all in series. When the switch has been closed for a long time, the current flowing will (ultimately) have a value of:

$$I = \frac{E}{R}$$

But if we assume that the resistance of the inductor is very small, then the current cannot reach this value for the following reason: if the current through the inductor changed instantaneously from zero to I, then the flux would also rise instantaneously to its final value.

However, we know that the voltage required to achieve this is given by the rate of change of flux. An instantaneous change of flux represents an infinitely large rate of change of flux, and this would require an infinitely large voltage. Consequently the flux must rise more slowly, as it is limited by the available voltage. *Figure 6* is a graph that shows the rise in current through the inductor. (It is, of course, the same shape as a graph that would show the rise in flux.)

The time constant, $\tau$, is defined as the time which would be taken for the current to reach its final value if it continued to rise in a linear manner. $\tau$ is given by:

$$\tau = \frac{L}{R}$$

$\tau$ is more precisely defined as the time taken for the current to rise to a value of 0.63 I.

**Decay of flux in an inductor**

We can use a similar argument to determine the way in which the flux or the current that exists in an inductor decays to zero, when the supply battery is disconnected and the resistor is connected directly across the inductor. □

ELECTRICAL TECHNOLOGY

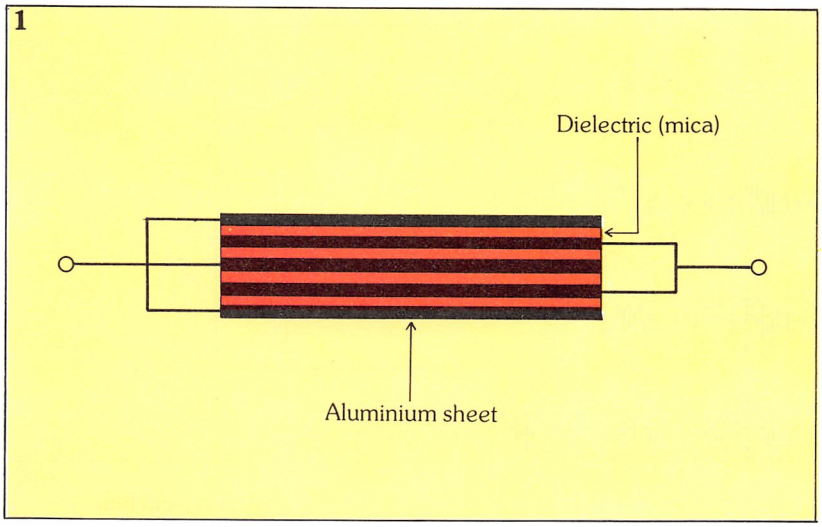
# Calculations of capacitors and inductors

In an earlier Basic Theory Refresher we found that the capacitance, C, between two metal plates of area A $m^2$ placed 1 m apart and filled with a dielectric of relative permeability $\varepsilon_r$ is given by:

$$C = \frac{\varepsilon_o \varepsilon_r A}{l}$$

This equation enabled the calculation of the capacitance of a very simple capacitor. We'll now take a look at three complex capacitors which are frequently used in electronic circuits.

*Figure 1* illustrates a parallel plate capacitor made from five thin sheets of aluminium separated by four layers of mica.




**1. A parallel plate capacitor.**

The mica is 0.3 mm thick and the area of each plate is 1.5 $cm^2$. As you can see, alternate aluminium plates are connected together to form the capacitor's two terminals. Mica has a relative permittivity of 6.0. With this information we can now find the capacitance:

$$C = \frac{8.85 \times 10^{-12} \times 6.0 \times 1.5 \times 10^{-4}}{0.3 \times 10^{-3}}$$

$$= 2.65 \times 10^{-11}\ F$$
$$= 26.5\ pF$$

Manufacturers need to state the maximum voltage that each type of device can withstand without breaking down; this **dielectric strength** can be defined as the maximum voltage per unit distance that a material can withstand. As the dielectric strength of mica is $5 \times 10^4\ Vmm^{-1}$, the maximum voltage for our capacitor can thus be found as follows:

$$\text{Maximum voltage} = 5 \times 10^4 \times 0.3$$
$$= 1.5 \times 10^4\ V$$
$$= 15\ kV$$

In practice, however, this capacitor would probably be rated for operation at a maximum voltage of 1 kV.

A typical film capacitor, made from two layers of metal film and two layers of polypropylene film, interleaved and rolled up to form a cylinder, is shown in *figure 2*. The area of the original sandwich is 20 $cm^2$; the dielectric is 0.05 mm thick and has a relative permittivity of 3.5. When the sandwich is rolled up, we effectively have two capacitors in parallel, because both sides of the metal foil are used. Thus, the total capacitance can be found:

$$C = \frac{2 \times 3.5 \times 8.85 \times 10^{-12} \times 20 \times 10^{-4}}{0.05 \times 10^{-3}}$$

$$= 2.48 \times 10^{-9}\ F$$
$$= 2.48\ nF$$

As a final example, let's consider a ceramic capacitor constructed using a wafer of barium tilanate as a dielectric. This wafer is 6.5 mm long, 7.5 mm wide and 0.2 mm thick with each side coated with metal to form the two plates of the capacitor. The relative permittivity of the ceramic wafer is 2000, so we can calculate the capacitance to be:

$$C = \frac{2000 \times 8.85 \times 10^{-12} \times 6.5 \times 10^{-3} \times 7.5 \times 10^{-3}}{0.2 \times 10^{-3}}$$

$$= 4.3 \times 10^{-9}\ F$$
$$= 4.3\ nF$$

## Inductors

We'll now turn our attention to a determination of inductance; the type of inductors to be discussed are used in both power circuits and communications systems. Iron is often used to make inductor cores as it has a permeability much greater than that of air. Power supplies and audio systems usually employ inductors with laminated steel cores, while radio frequency circuits and some audio applications demand the use of ferrite cored inductors.

First, we'll determine the inductance of a coil comprising 200 turns of wire wound onto a laminated ring of stalloy stampings. The mean diameter of the ring is 40 cm and the cross-

sectional area is 5 $cm^2$; we will assume that the magnetising current reverses from 1 A to −1 A.

At its maximum current the magnetic field strength is given by:

$$H = \frac{NI}{l}$$

$$= \frac{200 \times 1}{40 \times 10^{-2}}$$

$$= 500\ Atm^{-1}$$

The relative permeability of stalloy is 1750, thus the reluctance is:

$$S = \frac{l}{\mu_o \mu_r A}$$

$$= \frac{40 \times 10^{-2}}{1.26 \times 10^{-6} \times 1750 \times 5 \times 10^{-4}}$$

$$= 3.63 \times 10^5\ AtWb^{-1}$$

The inductance is given by:

$$\frac{N^2}{S} = \frac{200^2}{3.6 \times 10^5}$$

$$= 0.11\ H$$

If we now consider a current change from 8 A to −8 A then:

$$H = \frac{200 \times 8}{40 \times 10^{-2}}$$

$$= 4000\ Atm^{-1}$$

At this magnetising field strength, the relative permeability is known to be 300. Thus, the reluctance is:

$$S = \frac{40 \times 10^{-2}}{1.26 \times 10^{-6} \times 300 \times 5 \times 10^{-4}}$$

$$= 2.12 \times 10^6\ Atm^{-1}$$

and hence the inductance is:

$$L = \frac{200^2}{2.12 \times 10^6}$$

$$= 0.019\ H$$

which shows us that the inductance of an iron cored inductor varies with the value of the current which is being switched.

A coil of 150 turns wound onto a ferrite core is shown in *figure 3*. The effective cross-section of the core is 52.5 $mm^2$ and its effective length is 57.5 mm. The relative permeability of the ferrite used is 1800. The inductance:

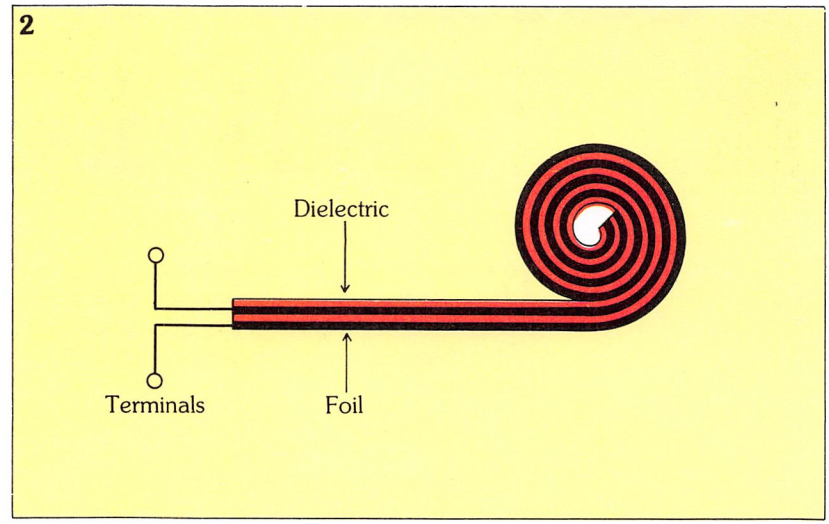
$$L = \frac{\mu_o \mu_r N^2 A}{l}$$

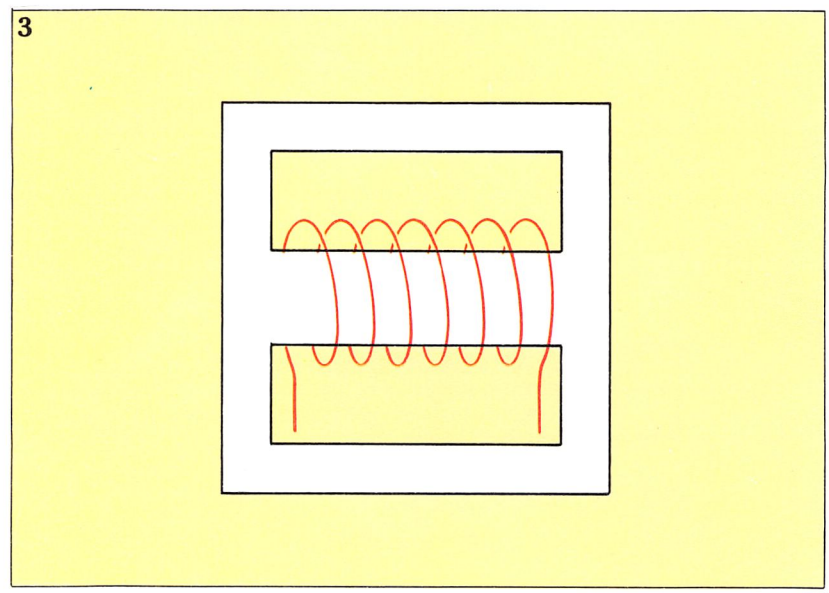
$$= \frac{1.26 \times 10^{-6} \times 1800 \times 150^2 \times 52 \times 10^{-6}}{57.5 \times 10^{-3}}$$

$$= 0.046\ H$$

$$= 46\ mH$$

□




**2. A typical film capacitor.**

**3. A coil of 150 turns** wound onto a ferrite core.

18

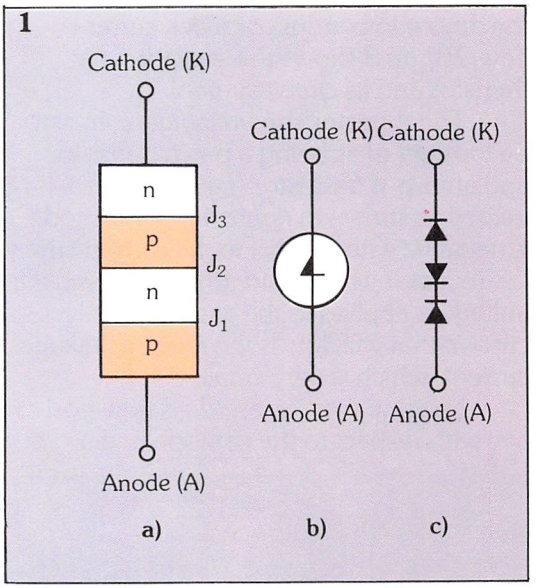
SOLID STATE ELECTRONICS

# p-n-p-n devices - thyristors and triacs

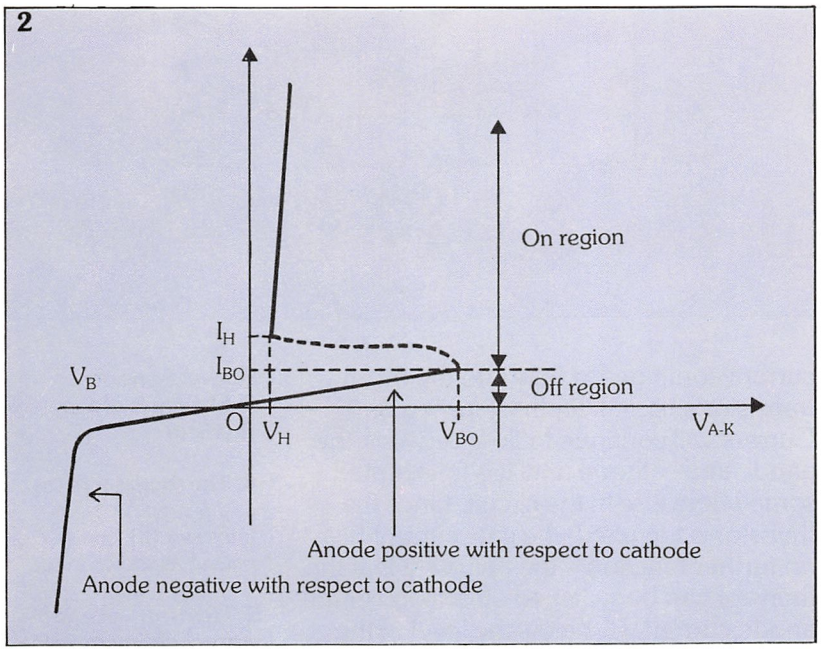
## Multilayer semiconductors

In this chapter we shall look at the operation and applications of two multilayer semiconductor devices – the **thyristor** and the **triac**. These devices may be considered in the simplest sense as controlled diodes, i.e. diodes that only start conducting when control pulses are applied to them. The thyristor, in fact, is often referred to as a **silicon controlled rectifier** (SCR).

Operation of thyristors and triacs is best understood if we first consider a basic multilayer device made from four alternate p and n-type layers of semiconductor material (*figure 1a*). An electrical circuit symbol we may use for such a device is shown in *figure 1b*. These components can be thought of as three diodes arranged in series, as *figure 1c* demonstrates. If a voltage is applied across the device, such that the cathode is positive with respect to the anode, the two outside junctions ($J_1$ and $J_3$) are reverse biased while the centre p-n junction is forward biased. The device operates like a single reverse biased diode, only conducting to any extent when the breakdown voltage of p-n junctions $J_1$ and $J_3$ has been exceeded – this is shown in the characteristic curve of *figure 2* as the lower curve, i.e. from 0 V to $V_B$.




**1. (a) Structure of a basic multilayer device; (b)** its circuit symbol; **(c)** this device can be thought of as 3 diodes connected in series.




**2. Characteristic curves** for a multilayer device.

If the applied anode voltage is positive with respect to the cathode, junctions $J_1$ and $J_3$ conduct but the junction $J_2$ is now reverse biased. As the voltage is increased (from 0 V to $V_{BO}$ in *figure 2*), the reverse current increases up to the point where the breakdown voltage of junction $J_3$ is exceeded ($V_{BO}$). At this point, the current through the device abruptly rises while the voltage across it falls – the broken line portion of the curve. The four-layer diode has switched from its off state to its on state.

The voltage at which the diode is switched on is called the **trigger voltage** or **breakover voltage**, $V_{BO}$. To bring the diode back to the off state once conducting, the current intensity must drop – even for a brief moment – below the level of the **holding current**, $I_H$.

# Thyristors

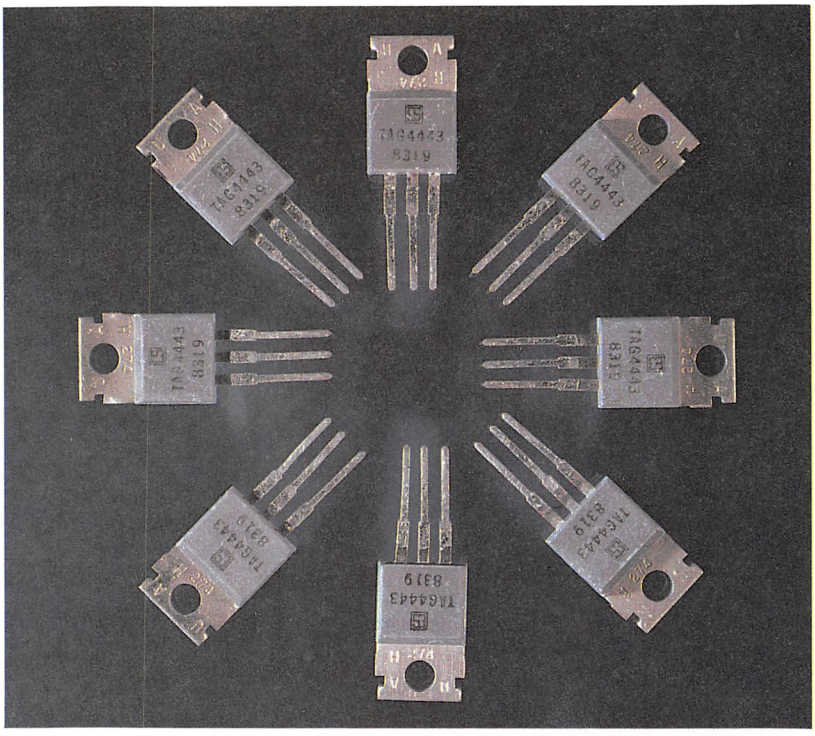
Thyristors are four layer semiconductor devices which can be made to switch currents of up to thousands of amps. They differ from the basic four layer diodes we have just seen in that they have *three* terminals: the **anode**, **cathode** and **gate**. The gate is a control terminal that enables the device to conduct or block current flow. *Figure 3* shows the structure of a thyristor and its circuit symbol.

The thyristor shown in *figure 4a* can be thought of as being a p-n-p transistor and an n-p-n transistor (*figure 4b*) connected as shown in *figure 4c*. If the anode is negative with respect to the cathode then the thyristor will remain turned off even if a voltage is applied to the gate. The only current that will flow is the reverse leakage current, which is very small.

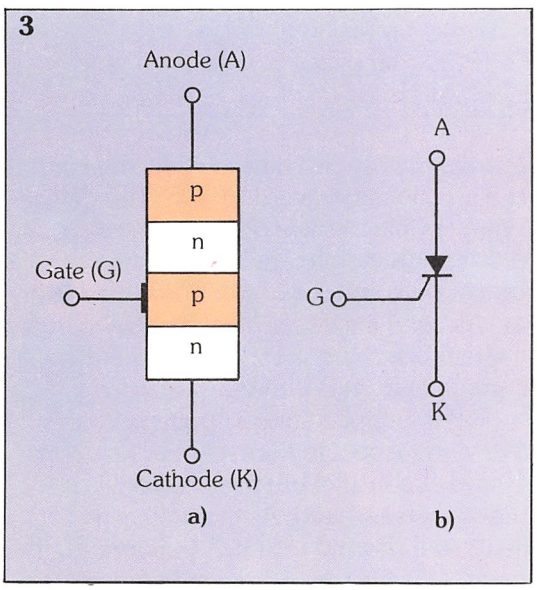
Imagine that the anode is now positive with respect to the cathode and a positive voltage is applied to the gate. The two outer junctions will become forward biased while the central junction remains reverse biased. However, the positive voltage applied to the gate creates a base current in the n-p-n transistor and turns it on. The n-p-n transistor's collector current $I_{C2}$, is the base current from the p-n-p transistor, so the p-n-p transistor turns on, as well.

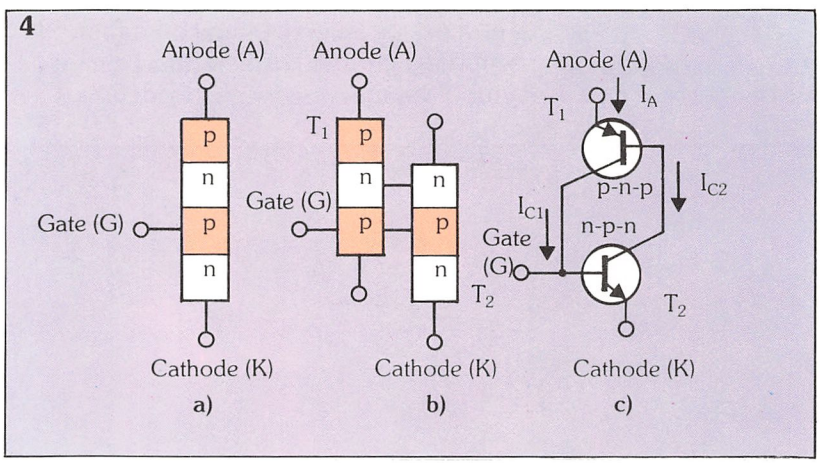
Even if the gate current is now removed, the p-n-p transistor's collector current continues to flow into the n-p-n transistor's base, keeping it turned on. Current will continue to flow between the anode and cathode until it is interrupted somewhere else in the circuit. Once the thyristor is triggered, the gate current has no further function – the only way that the thyristor can be switched off is to lower the anode current, $I_A$, below the level of the holding current or to simply switch the circuit off for an instant.

*Figure 5* shows thyristor's voltage-current characteristic. We can see that it behaves like a reverse biased diode for positive voltage and current levels (first quadrant) until the trigger voltage is ex-



Neville Miles



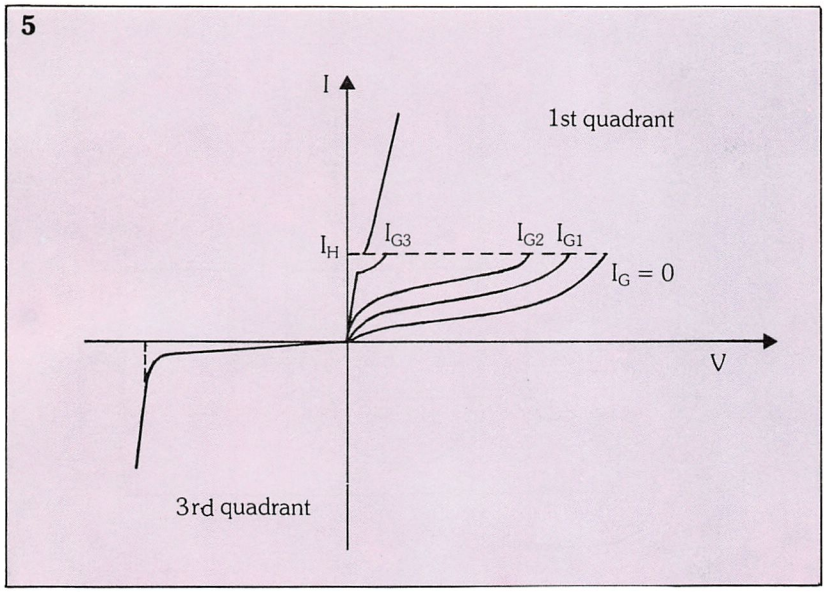

**Above:** thyristors.
**3. Structure of a thyristor.**

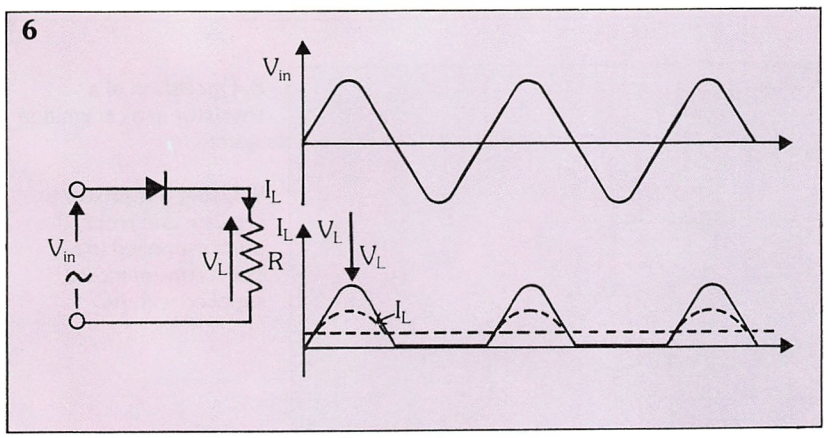
**4. The thyristor in (a)** as a p-n-p and n-p-n transistor **(b)**, connected as shown in **(c)**.

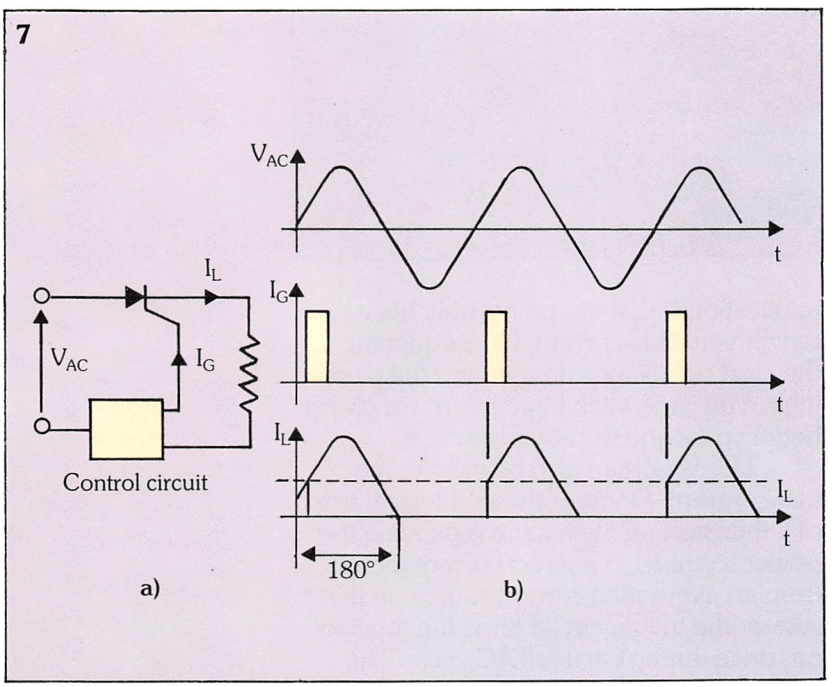
**5. Thyristor voltage current characteristic.**

**6. Basic half-wave rectifying circuit.**

**7. Replacing the diode from** *figure 6* with a thyristor produces this circuit.




ceeded. Once this happens the characteristic is similar to a forward biased diode.

The trigger voltage varies according to the gate current. The higher the gate current, the lower the voltage needed to trigger the thyristor. This gives us a family of characteristics which depend on the value of the gate current, $I_G$. Temperature variations can, however, affect thyristor triggering, with low temperatures making triggering harder.

**Thyristor applications**

It is often necessary to be able to control the current or voltage supplied to electrical equipment: thermostats control temperatures; dimmers adjust the intensity of lighting; and motor speed also needs to be regulated. In the past, these types of applications were handled by variable or fixed transformers, or by a variable or fixed control resistance – both costly and inefficient control methods.

Thyristors have allowed a different type of control circuitry to be devised, that is cheaper, simpler and more efficient.

When we looked at the operation of power supply circuitry in *Solid State Electronics 8*, we investigated the basic half-wave rectifying circuit (*figure 6*). The diode – being a unidirectional device – prevented the passage of the negative half-wave. This gave a pulsing output voltage which could be further filtered and levelled.

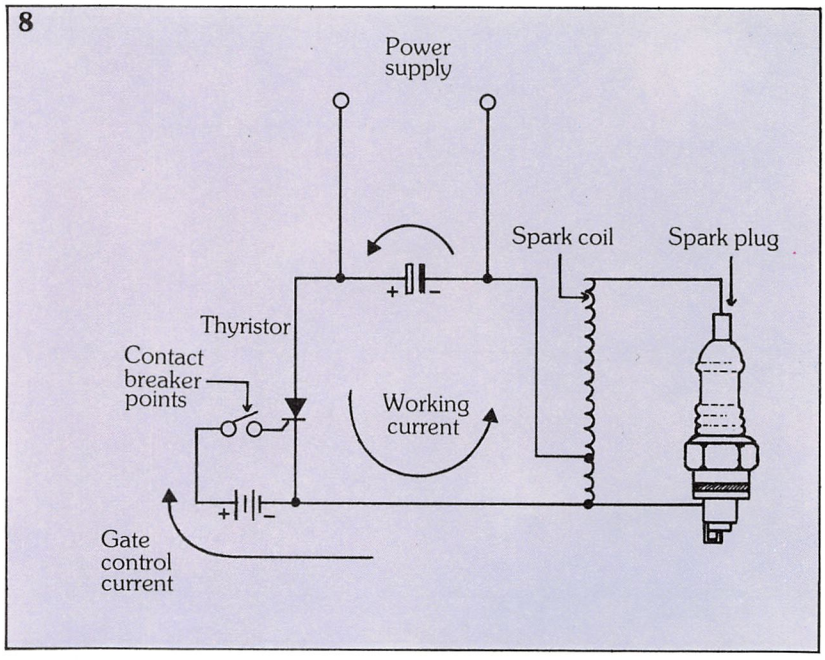
Now if we replace the diode with a thyristor, we will get the circuit shown in *figure 7a*. The basic difference here is that the thyristor can be externally controlled. This means that the half-wave rectified sinusoidal current output is only sent after the thyristor has received a command pulse from the control circuit.

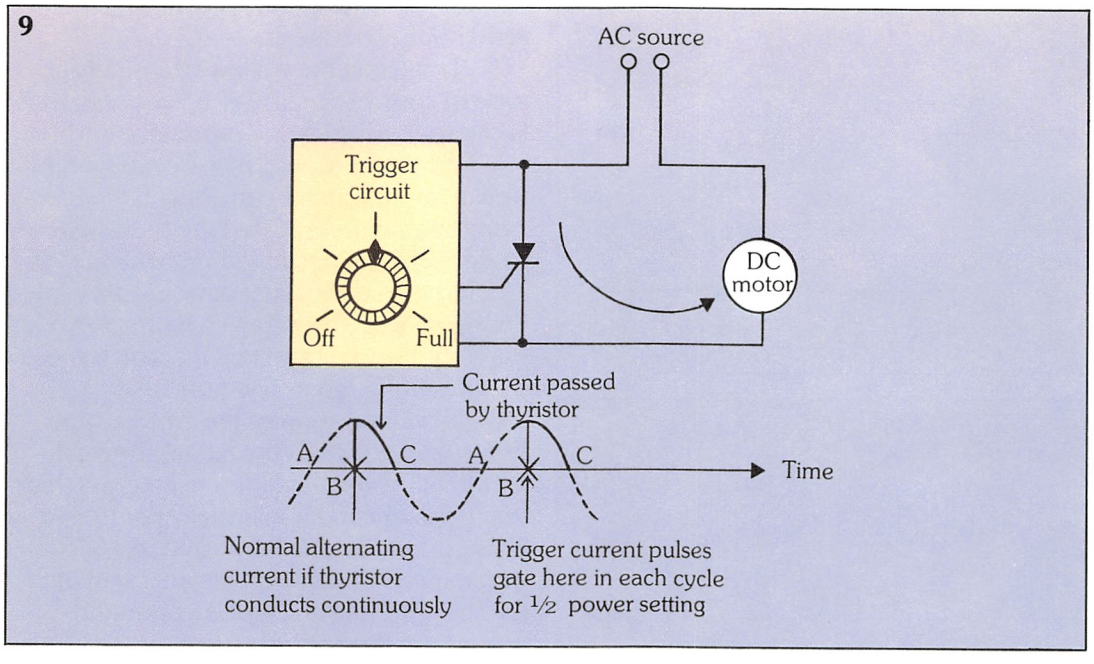
By choosing the moment at which the thyristor is triggered during a half-cycle, we can obtain an output voltage with a previously chosen average value. This average value can be changed by varying the moment at which the ciruit is triggered. The three graphs in *figure 7b* show the input voltage waveform, a set of control pulses and the circuit's consequent output waveform.

Generally speaking, the control circuits that decide the triggering of thyristors are rather complex. The control

circuits may actually employ other devices from the thyristor family, and we shall look at some of these components in a later chapter.

*Figure 8* shows the operating principle of a possible thyristor application – an electronic car ignition system. The power supply circuit stores up electrical charge in the capacitor while the thyristor is turned off. When the spark plug fires, the contact breaker points close momentarily (unlike a conventional ignition system where the points *open* when the spark plug is to fire) so that a small control current reaches the gate of the thyristor. This triggers the thyristor to the on state, discharging the capacitor through the few primary turns at the lower end of the spark coil. This sudden high current surge induces a very high voltage in the secondary turns of the spark coil, which in turn generates the spark across the plug's spark gap. The spark is only present for a fraction of a second, and once the capacitor charge is discharged the working current through the thyristor stops, turning the thyristor off and allowing the capacitor to recharge.




**8. Operation of a thyristor** in a car ignition system.

**9. Using a thyristor** to regulate and rectify the power supplied to a direct current motor supplied with AC.

Ordinary non-electronic ignition systems use the contact breaker points to switch the high working current directly. As a result, the points wear out quite rapidly causing poor engine performance. The great advantage of the thyristor in this application is that the points only have to switch very little current. Consequently they last significantly longer and the working circuit can switch higher currents giving better spark and improved ignition.

Thyristors can also be used to regulate current. *Figure 9* shows a typical use of a thyristor: rectifying and regulating the power supplied to a direct current motor from an alternating current source. In this system the trigger circuit turns the thyristor on, once during each full AC cycle. The

point during the cycle where the trigger pulse occurs is set by the manual control knob. By altering the trigger point, different amounts of power will be passed. The graph in *figure 9* shows the AC source by the broken line and the rectified thyristor output by the solid line.

In this example, the thyristor output is set to half power, because the trigger pulse occurs halfway through the positive waveform. At point A in each cycle, current tries to pass forward (from anode to cathode), but is blocked because the thyristor isn't turned on. At the phase of the cycle indicated by point B, the trigger pulse turns on the thyristor and current flows through it. At point C, the forward current has decreased to zero so the thyristor turns off again. The trigger pulse timing can be varied to occur at any point in the cycle between points A and C, thereby varying the average current supplied to the motor across the range from zero to complete forward half-waves.

This method of power regulation is called **phase control**. Its advantage is that virtually no power is wasted (dissipated) by being converted to heat. Heat is generated when we regulate current, not when we switch it. When we control current through some sort of regulating control element like a transistor or a rheostat, a voltage drop occurs across the regulating element, causing a current to flow – and this generates heat. But the thyristor is always either fully on or off: it never controls current partially.

## Triacs – theory and operation

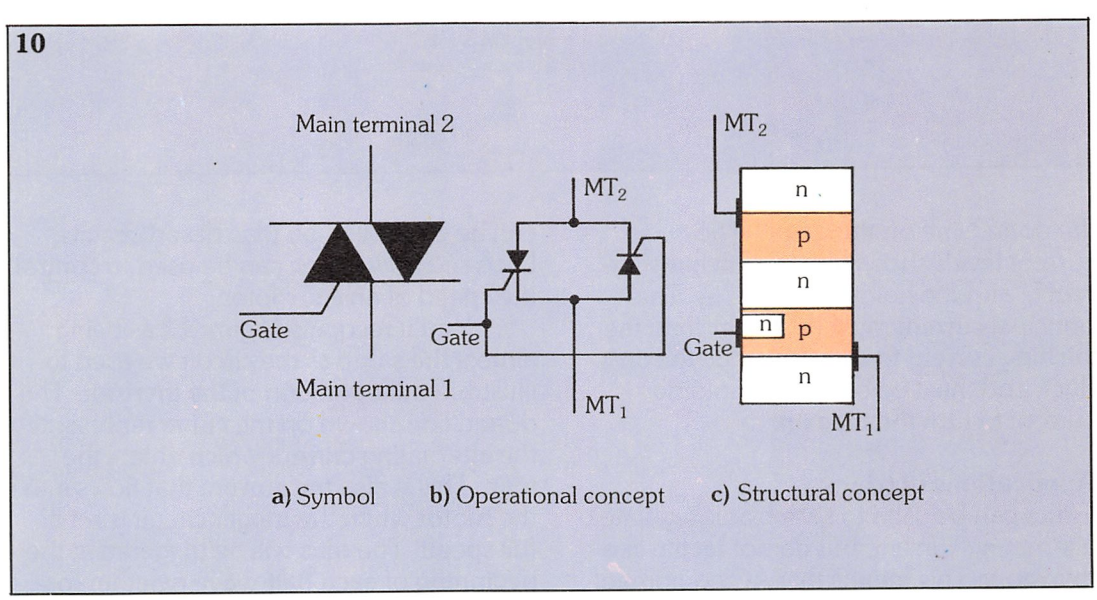
The triac (**tri**ode **a**lternating **c**urrent) is a semiconductor device with three terminals. It behaves very much like a pair of thyristors connected head to tail. The symbol for a triac shown in *figure 10a* reflects this. *Figure 10b* makes this clearer, the left-hand thyristor conducts conventional current in the downward direction and the right-hand thyristor conducts in the opposite direction. Notice that the gate terminal is common to both thyristors.

Picture the triac in an alternating current circuit – it normally blocks working current attempting to pass in either direction between the two main terminals ($MT_1$ and $MT_2$) but it can be triggered to conduct in both directions by a momentary positive or negative pulse applied to the gate. This is reflected in the triac's formal name, which is 'bidirectional triode thyristor'.

The triac semiconductor element has five layers and a small sixth region under part of the gate contact (*figure 10c*) permitting the triac to be triggered by gate current moving in either direction – not just into the gate as in the thyristor.

### Electrical characteristics of a triac

The graph in *figure 11* illustrates a triac's static characteristics – notice that as it conducts in both directions, the reverse and direct characteristics are symmetrical.

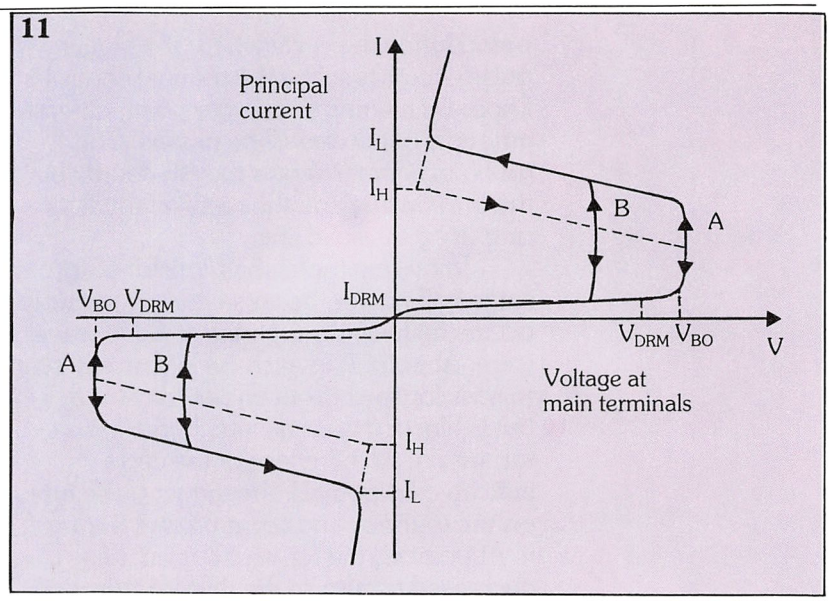


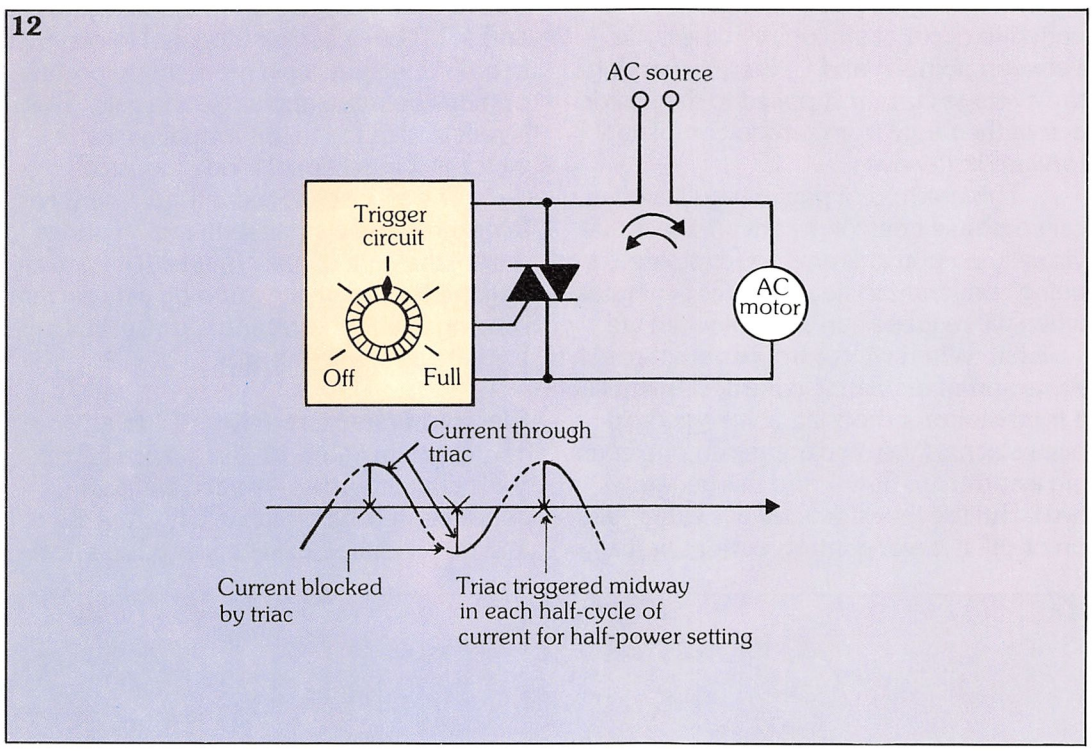

**10. A triac: (a)** symbol; **(b)** operational concept; and **(c)** structure.

It is also worth noting that the triac's terminals are called **main terminals** ($MT_1$ and $MT_2$) because it conducts in both directions and therefore cannot have a cathode and an anode.

When a triac is conducting, the current that flows between the main terminals is known as the **principal current**. Curve A in *figure 11* is the I-V characteristic of a triac with zero gate current: both directions of applied voltage are shown. The principal current is small and reasonably constant until the breakover voltage ($V_{BO}$) is reached. Curve B, on the other hand, shows the I-V curve when there is a gate current present.

As you can see, there is a hysteresis effect along the current axis indicated by




**11. Characteristic curves of a triac.**

**12. Using a triac to switch** and regulate current in an AC motor.

the dotted line on the graph. The two current levels shown are the latching current $I_L$ and the holding current $I_H$. The principal current must be higher than the latching current to allow the triac to conduct, and must be below the holding current to turn the triac off.

### Applications of triacs

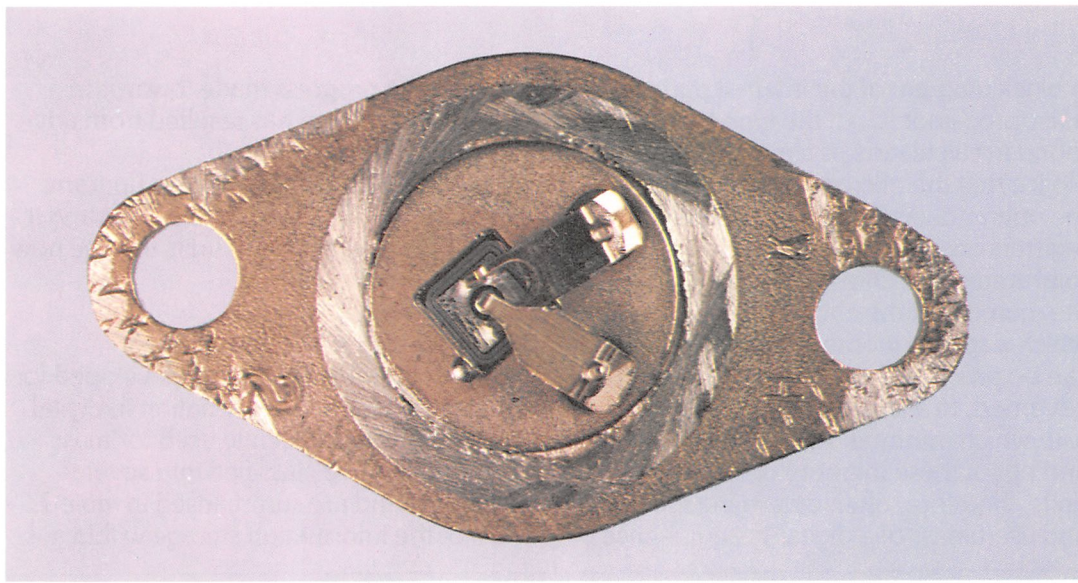
Triacs can be used to switch and regulate alternating current, but do not rectify like thyristors. This means that AC equipment can be controlled by triac based circuits. Let's see how a triac can be used to control the speed of an AC motor.

You'll recognise *figure 12* as being almost the same as the circuit we used to illustrate the operation of the thyristor. The dotted line shown on the curve represents the alternating current which enters the triac. This is also the current that flows into the motor when the trigger circuit is set at full speed. The triac will be triggered at the beginning of each half-wave whether for-

ward or reverse biased.

If we set the knob to half speed, then the control circuit delays triggering the triac until the middle of each half-wave. This is shown by the solid curves which represent the triac's output. This resulting waveform is still AC, but since, on average, less current passes each way, the motor is slowed down. Thus the triac can *regulate* AC by means of phase control switching. This is employed in speed controls for electrical machinery and light dimmers.

**Right:** triac with the top cut away showing connections.



Neville Miles

## Glossary

| | |
|---|---|
| **gate** | control terminal of triacs and thyristors |
| **holding current** | minimum level of current passing through thyristors or triac which will hold it in the on state |
| **main terminal** | thyristors only conduct in one direction, so their main terminals are named anode and cathode. Triacs, on the other hand, conduct in both directions, so their principal connections are simply called main terminals ($MT_1$ and $MT_2$) |
| **phase control** | operation that sends a trigger pulse to a thyristor or triac in each half-wave, causing the device to switch on for the remainder of the half-wave. By altering the point in each half-wave at which the trigger pulse is sent, a variable output is obtained |
| **thyristor** | controllable p-n semiconductor device that rectifies and switches AC giving a DC output. Can be used to regulate and control DC equipment |
| **triac** | controllable p-n-p-n semiconductor device that switches and regulates AC, giving an AC output. Operates like two thyristors connected head to tail in parallel , enabling it to conduct current in both directions |
| **trigger voltage** | the level of voltage applied momentarily to the gate of a thyristor or triac, that causes the device to switch on. Level varies according to the gate current. Also known as breakover voltage |

16

DIGITAL ELECTRONICS

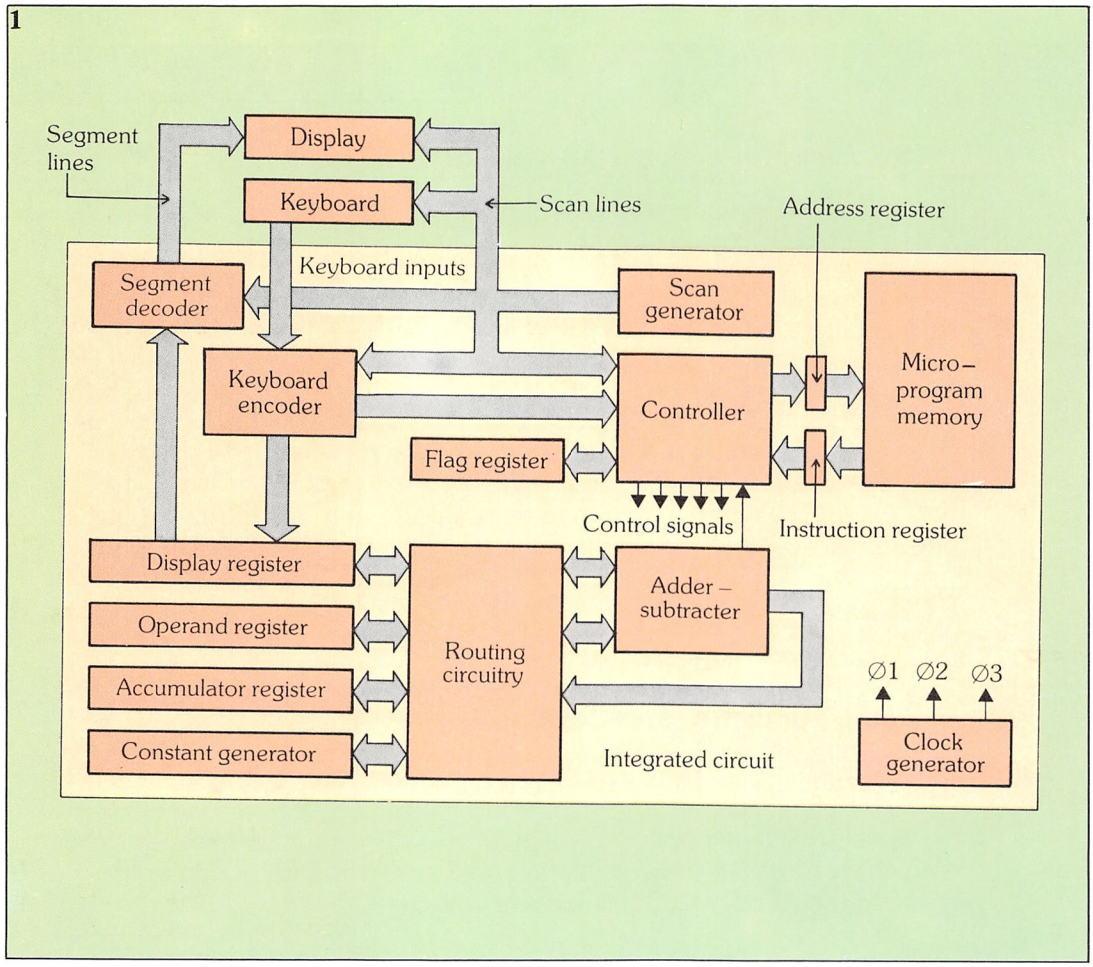
# Memories in digital systems

A block diagram of the internal parts of a microprocessor IC, of the type typically found in calculators, is shown in *figure 1*. Notice that the microprogram memory and the operand, display and accumulator registers occupy a significant portion of the total area of the chip: these areas comprise the memory of the chip. The speed with which a system works (i.e. how much data can be processed in a given time) may be governed, to a large extent, by the speed with which information can be moved into and out of these memory units. Memory units, therefore, often determine the cost and usefulness of a digital system, hence much of the progress made towards increased integration has resulted from advances in their design.

A memory unit is one step up from simple storage devices like flip-flops and it is to a discussion of these units that we now turn.

## How memories are classified

Many techniques have been developed for the storage of bits of information in digital systems – those presently used for mass memories can be classified into several categories and are summarised in *table 1*. Most of the information storage within a




**1. Block diagram** of the internal parts of a microprocessor IC.

microprocessor IC is achieved with charge storage RAM, or ROM.

Notice in *table 1* that memory types are not only classified according to storage method but also by method of access. The access method of a memory is simply the order or sequence in which stored information is retrieved. In a **serial access** memory the stored information is available for reading only in a certain order, usually the same order in which it was written. In a **direct access** memory, on the other hand, information stored may be read in any order and is consequently significantly faster than serial access memory. (Direct access memories will be discussed in *Digital Electronics 17.*)

The Research House/Keen Computers

**Above:** CAD/CAM terminals, forming part of an integrated computer system for manufacturing. Over 1000 terminals can be linked via a network.

## Serial access memory

Although the punched tape and card methods of data storage are largely obsolete, they do provide good examples of serial access memory. Punched tape consists of a length of paper in which holes are punched to represent logic 1; the lack of a hole represents logic 0. The holes and spaces are arranged in columns across the width of the tape constituting binary words. The words are aligned vertically along the length of the tape.

Data is written to the tape by a punch head which writes (or punches) a word at a time; data is read a word at a time by another head where small light sources on one side shine through the holes in the tape to be read by an equal number of photosensitive devices on the other.

This kind of memory is extremely slow and cumbersome. Access to one particular data word requires serially reading all words up to that word on the tape.

Magnetic tapes, disks and drums are also examples of serial access memory and were discussed in *Basic Computer Science 4.* Bits are recorded on the surface film of magnetic material by applying a magnetic field in one of two directions; the bits are stored in lines and columns in the form of tiny magnetised dots.

The types of memory discussed so far (and shown on the right of *table 1*) have enormous memory capacity at a very small cost per bit, however they are slow, i.e. their access time may be very long. For this

Table 1
**Classification of memory types according to access and storage methods**

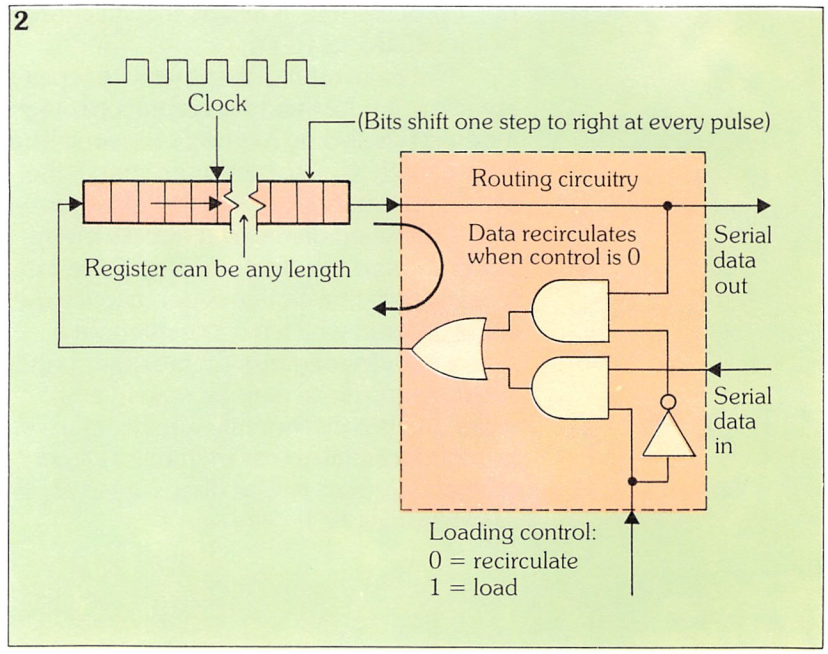
| Storage Method / Method of Access | Using integrated semiconductor storage elements | | | Using tiny permanent magnetic fields | | | Using punched holes | |
|---|---|---|---|---|---|---|---|---|
| | Flip-flops | Charge storage | Permanent connections | Fixed medium | Moving medium | | | |
| Serial Access | Static shift registers | Dynamic shift register | | Bubble | | Tape | Punched paper tape | Punched cards |
| Direct or Random Access | Static RAM | Dynamic RAM | ROMs | Cores | Disks and drums | | | |

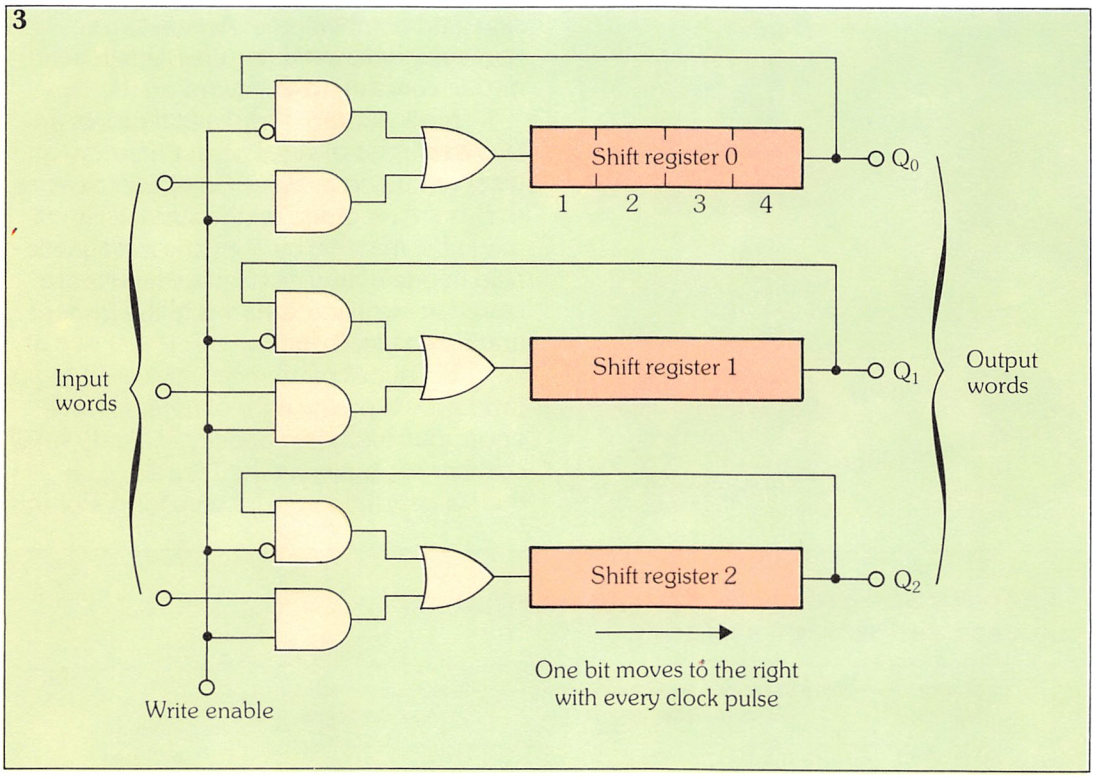
reason they are not used as main or central memory within digital systems and are considered as auxiliary, or peripheral memory.

### Serial access memories using shift registers

As we have seen, a shift register comprises a string of 1-bit storage units with a common clock. At each clock pulse, the stored bits shift one step along the string. A shift register can be used as a memory unit by **recirculating** data, as shown in *figure 2*.

The bits in the shift register are continually shifted to the right by regular clock pulses, with no pauses. When the **loading control** input in *figure 4* is at 0, each bit that is shifted out at the right-hand side of the register, is shifted back in at the left via the **recirculation path**. In this way,




**2. Recirculating data** enables a shift register to be used as a memory unit.

**3. A serial access memory** 3-bits wide and holding 4 words.

the register outputs each bit, one at a time, over and over again. Recirculating shift registers can be made which store any number of bits.

A serial access memory 3-bits wide and holding four words is illustrated in *figure 3*. The three shift registers shown each have four stages, providing storage for four words; each register stores one of the three bits in each word. With every clock cycle (the clock terminal is not shown in *figure 3*) the bits in each register move forward by one place, so the words appear in sequence at the outputs $Q_0$, $Q_1$, and $Q_2$.

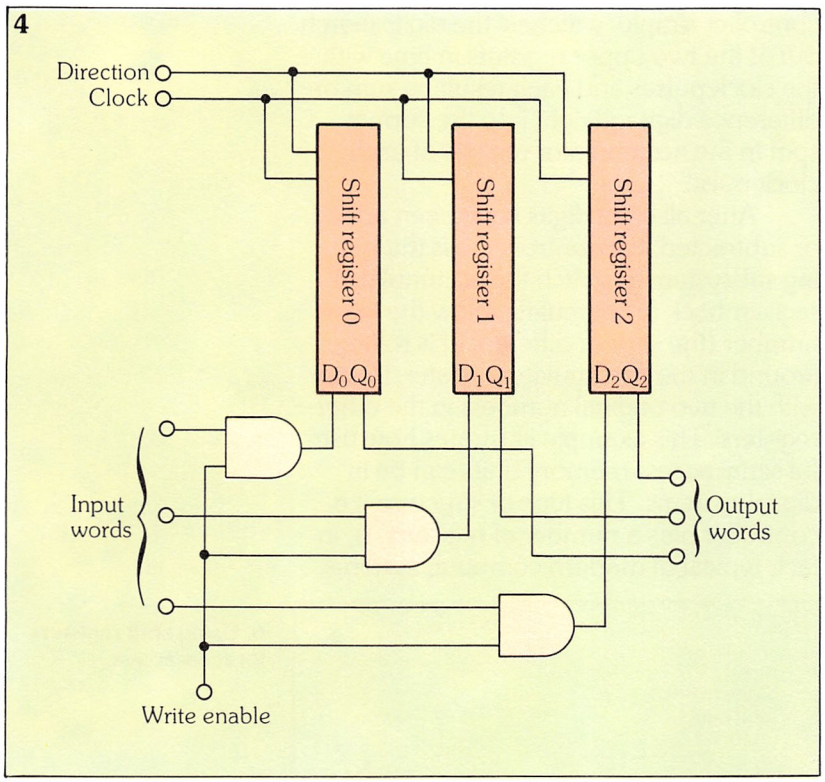
This type of memory is similar to paper tape memory with each end joined together, forming a continuous loop. If the **write enable line** is set to logic 1, the

existing contents of the shift registers may be replaced by new data appearing at the input. The contents of this type of memory appear at the output in the same order that they are placed into the memory, so this memory is a first in/first out system.

*Figure 4* illustrates another type of serial access memory also using three registers, each providing storage for 3-bit words. It differs from the previous example in that there are no recirculation paths and control input to the memory controls the shift register's direction.

When the memory is enabled, and the direction of data flow upwards, a sequence of input words can be stored. Each word enters the memory from the bottom and moves towards the upper stages, leaving space for the next word. The words can be thought of as being stacked one on top of another, with the last at the bottom of the stack. To read the words, the write enable line is set to logic 0 and the shift direction is reversed. With every clock cycle, the words then reappear at the output in the opposite order to which they were recorded. This, therefore, is a first in/last out memory system.

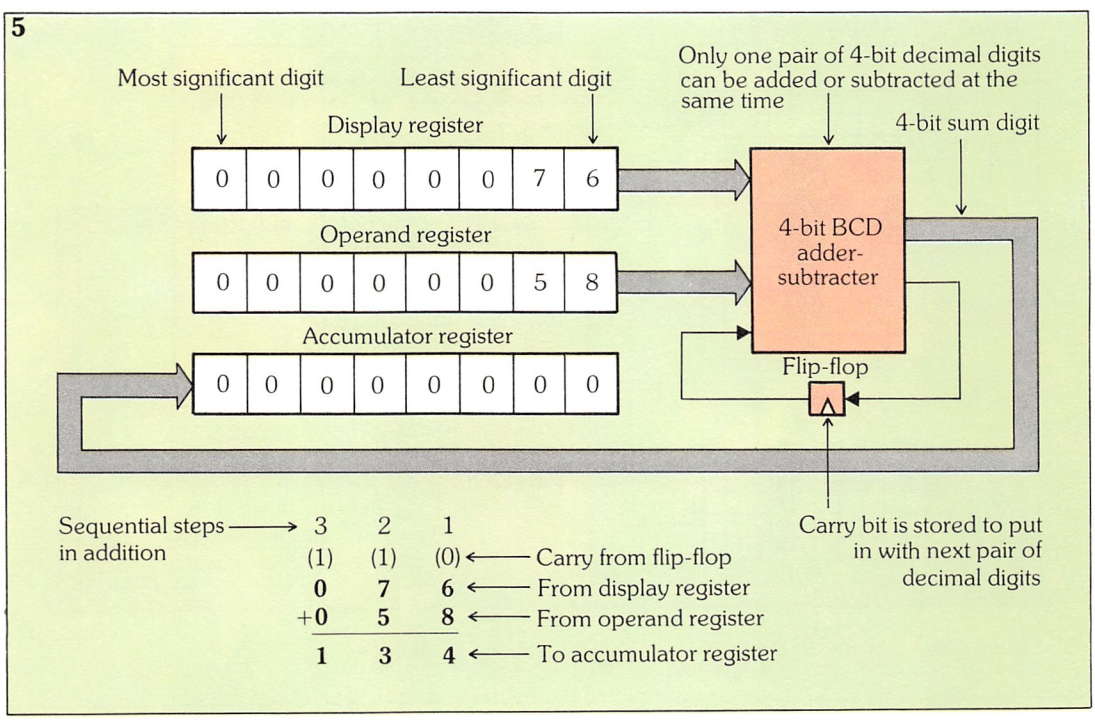
**4. Another example of serial access memory** – this time with no recirculation paths.




## How a calculator uses a shift register as serial access memory

*Figure 5* illustrates (in a simplified form) why serial access is useful in the three number registers of the calculator microprocessor IC of *figure 1*. It comes in handy when the two numbers in the display register and the operand register are being added or subtracted. The digits of each number go to the adder-subtractor one at a

**5. Serial access memory** is useful when two numbers in the display and operand registers of a calculator are to be added or subtracted.



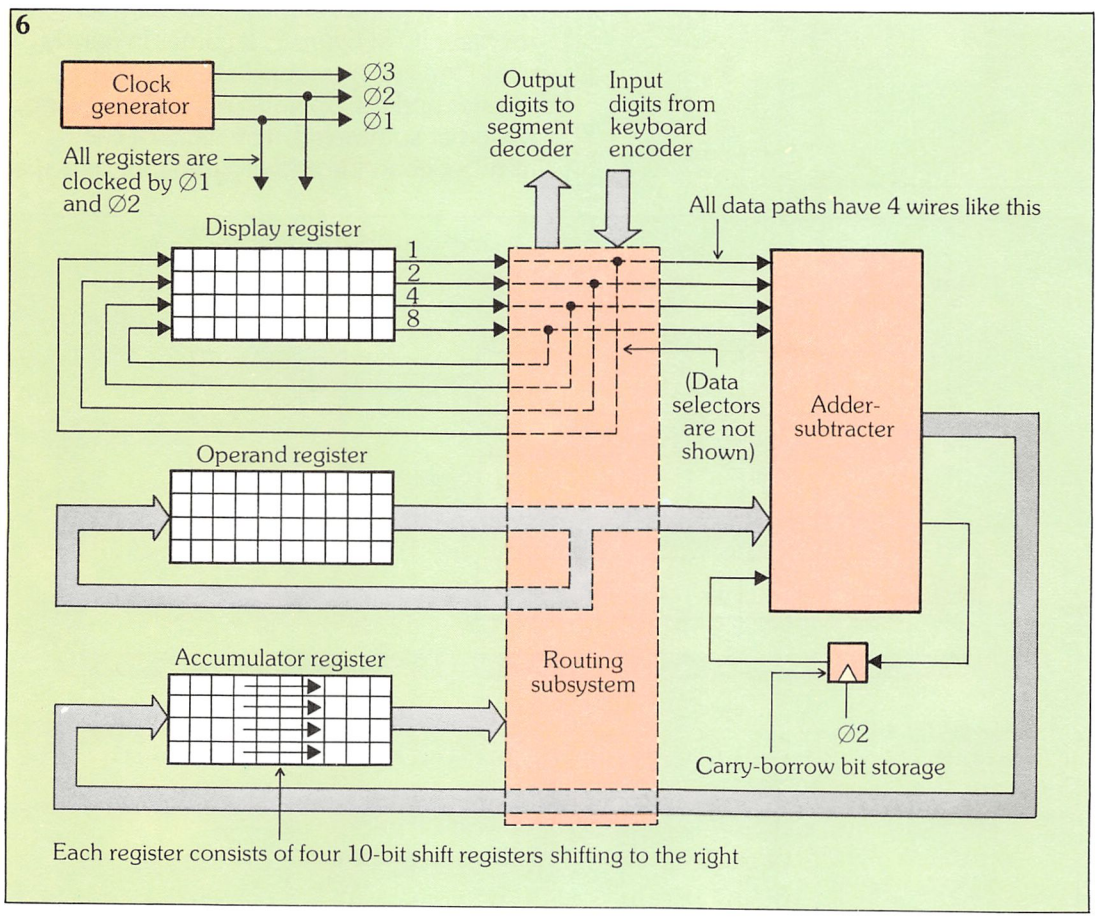
time in the order in which they are stored. First, the least significant digits (those on the far right) are added or subtracted, second, the next digits to the left, and so on. The adder-subtractor produces the appropriate sum or difference digit. These resulting digits must be loaded into the accumulator register in the same sequential order (least significant digit first).

As each pair of digits is added – say the 6 and 8 shown in *figure 6* – the adder-subtractor also produces a carry bit (which for 6 and 8 would be 1 along with a sum of 4). The carry bit is stored in a flip-flop and added in with the next most significant pair of digits (in this case 7 and 5). Similarly, a borrow bit is stored in the flip-flop during subtraction.

The use of shift registers for serial access is illustrated in *figure 6*: four recirculating shift registers are shown for each number register – one for each bit (8, 4, 2, 1,) in binary coded decimal digits. All twelve shift registers are clocked in step together, so that the bits for all three least significant digits appear at the outputs at the same time.

When the least significant digits appear at the outputs, the controller tells the adder-subtractor to begin adding (or subtracting); the controller then instructs the routing subsystem to route the adder-subtractor's output to the accumulator register input. The rest is automatic: the controller simply 'watches' the digits march out of the two upper registers in time with the clock pulses and each resulting sum or difference digit roll right into the correct spot in the accumulator register at each clock pulse.

After all eight digits have been added or subtracted, the controller tells the routing subsystem to switch the accumulator register back to recirculate. Now the new number (the sum or difference) is rolling around in the accumulator register, in step with the two original numbers in the other registers. This example illustrates how useful serial access memory units can be in digital systems. This type of structure, i.e. controller plus a number of registers, is, in fact, typical of modern computer systems.

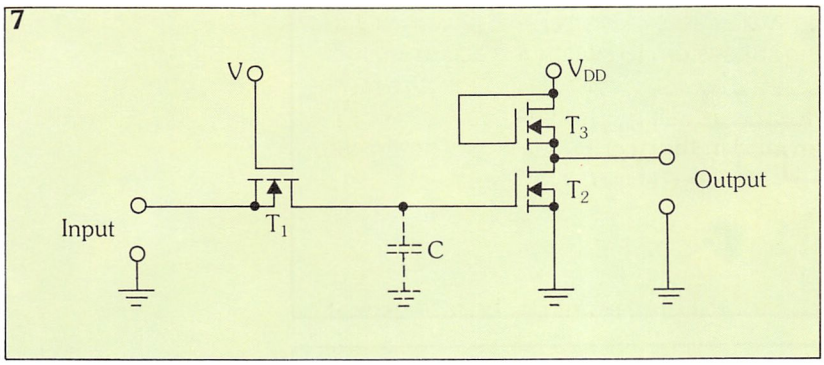



**6. Using shift registers** for serial access.

## MOS registers

Since MOS (metal oxide semiconductor) technology enables considerable integration it is used to manufacture most memory devices. MOS devices often use an electric charge method of operation known as **dynamic storage**. In *figure 7* a dynamic MOS inverter circuit capable of storing one bit for a small amount of time is shown. The circuit is constructed using NMOS transistors, i.e. n-channel enhancement type MOSFETs. Such transistors conduct when a logic 1 is applied to the gate.

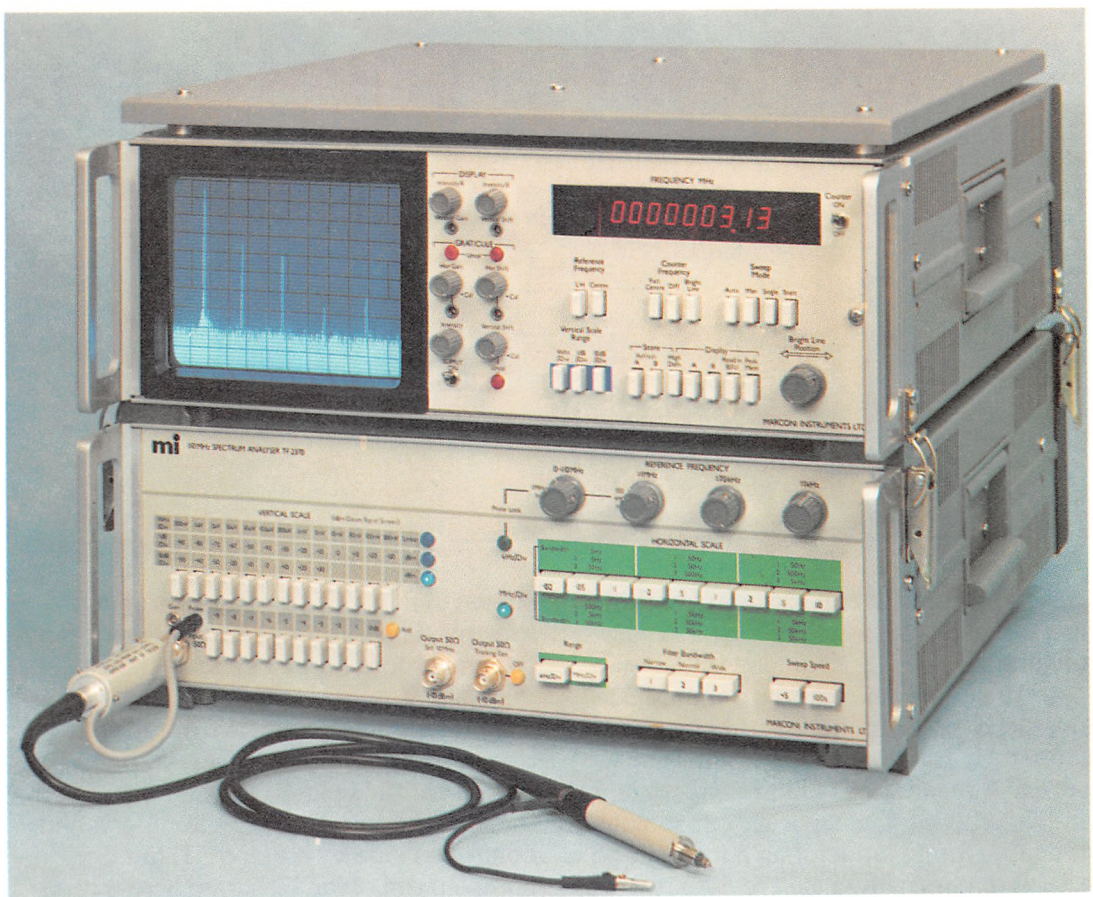
**7. A dynamic MOS inverter circuit** capable of storing one bit for a small amount of time.



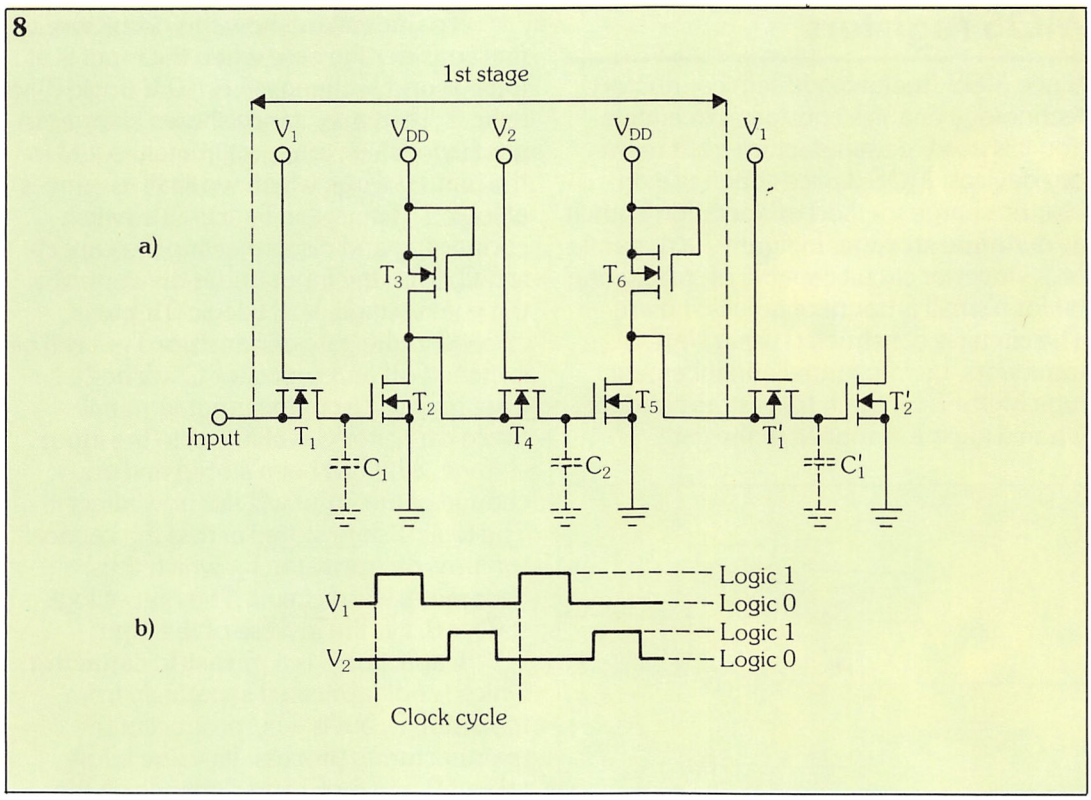

To understand how this circuit works, first consider the case when the input is at logic 1, and V changes from 0 V (logic 0) to logic 1. The capacitance shown charges to a voltage which is approximately equal to the input voltage which we shall assume is at logic 1. Transistor $T_1$ acts as a switch connecting and disconnecting the capacitor, C, from the input under the control of the gate voltage, V. If a logic 0 input is applied to the gate of transistor $T_1$, it will be switched off and capacitor C will be disconnected from the input terminal. Since capacitor C is charged to the input voltage, a bit has been stored and any change of the input will not now affect it. The logic 1 state stored across the capacitor turns on transistor $T_2$, which thus becomes a short circuit. The output logic state is 0, i.e. the inverse of the input.

Capacitor C is a **parasitic capacitor** which is not fabricated separately from transistor $T_2$ but is a by-product of the manufacturing process. Its value is low (about 0.5 to 5 pF) and depends on the type of transistor used. Often, the effects

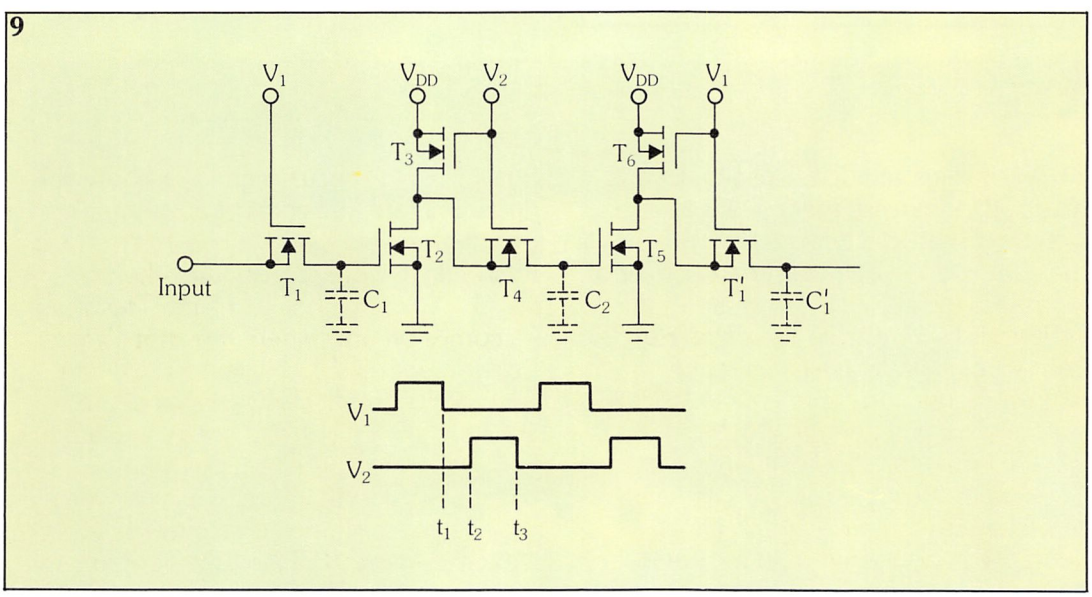
**Right:** spectrum analyser using digital storage to display information.




The Research House/Marconi Instruments




**8. (a) A dynamic MOS shift inverter; (b)** its timing diagram.



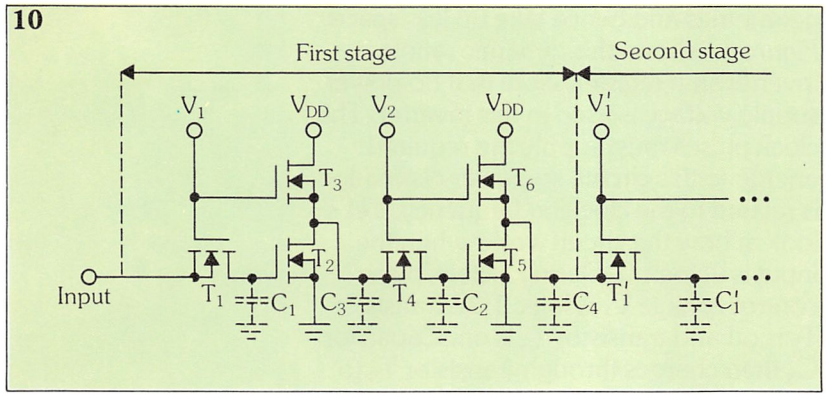

**9. Variation of circuit shown in** *figure 8* giving reduced power consumption.

caused by parasitic capacitors are undesirable but in this case they are used to advantage.

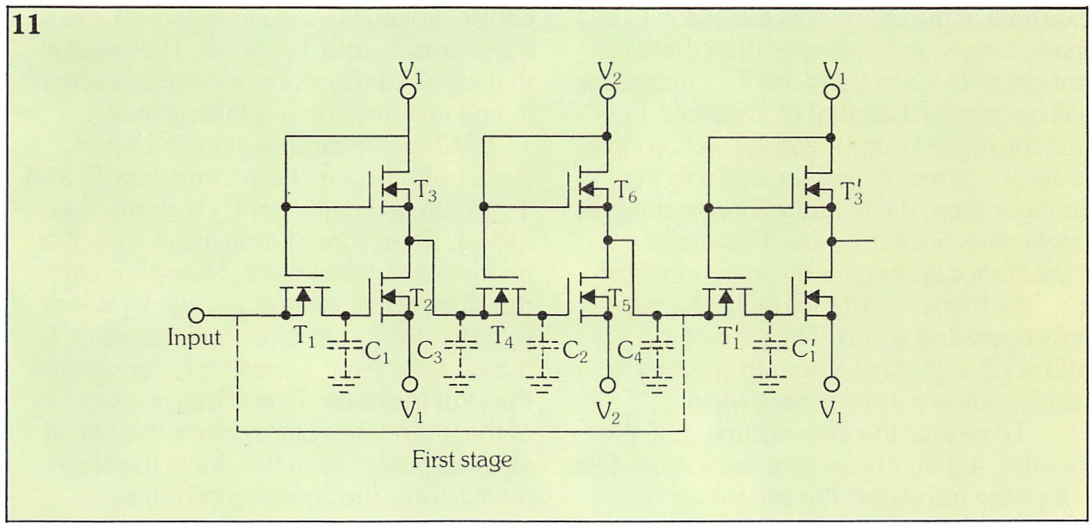
The information stored in this MOS inverter is, unfortunately, short lived because the capacitor discharges through **parasitic resistances** of the transistors and, in practice, the capacitor can maintain its charge for no longer than a millisecond.

A dynamic MOS shift register is shown in *figure 8a*. It is made from two inverter stages like the one in *figure 7*. The two transistors $T_1$ and $T_4$ need control voltages, $V_1$ and $V_2$, which are never logic 1 simultaneously, as shown in the timing diagram of *figure 8b*.

If control voltage $V_1$ becomes logic 1 the input is stored by capacitor $C_1$. When control voltage $V_2$ is at logic 1, the complement of the bit stored across capa-




**10. Another variation** of the circuit shown in *figure 8*.




**11. Dynamic ratioless inverter shift register.**

Now let's look at another important consideration, power consumption. When the voltage stored across capacitor $C_1$ is at logic 1, transistors $T_2$ and $T_3$ are both on, and current flows through them to earth, dissipating power. In the same way, if capacitor $C_1$ is discharged (logic 0) the complement of this will be found on capacitor $C_2$ (logic 1). This means that either the master or the slave is always dissipating some power.

Power consumption may be reduced by altering the circuit slightly, as shown in *figure 9*. Here, the gates of transistors $T_3$ and $T_6$ are directly controlled by the clock. Power is now only consumed when one of the clock signals becomes high. Transistor $T_3$ comes on only when transistor $T_4$ is on, just as transistor $T_6$ only conducts when transistor $T_1$ is on. Since neither of the clock signals is high for some of the time, you can see that less power will be consumed.

citor $C_1$ is transferred to capacitor $C_2$. At this point we can see that the effect of this circuit is similar to a master-slave flip-flop. The combination of transistors $T_1$, $T_2$ and $T_3$ can be called the master inverter, and transistors $T_4$, $T_5$ and $T_6$ the slave. Remember that a master-slave flip-flop needed complementary phased clock signals to operate, and note that this MOS circuit needs complementary phased gating voltages. This is because information must be present in the master section before it can be passed onto the slave section.

However, in contrast with a flip-flop, the rate at which the data is clocked through the MOS register must not fall below a minimum value. If the clock period is too long, the charge will leak out of the parasitic capacitors and the data will be lost. The clock period must therefore not last longer than a millisecond. This is why such memory circuits are known as **dynamic** – the stored data needs to be **refreshed**.

Another variation on this register is shown in *figure 10*. Here, the gates of transistors $T_3$ and $T_6$ are connected to the clock of the preceding part, rather than the following part. In this case, when control voltage $V_1$ is logic 1, the data input is transferred to capacitor, $C_1$, while its complement is stored across $C_3$. When control voltage $V_2$ is logic 1 the voltage across capacitor $C_3$ is transferred to capacitor $C_2$, and the complement is again transferred through the inverter, this time to capacitor $C_4$. The circuit contained within the broken lines is the first stage of a shift register.

**Other types of two-phase shift register**

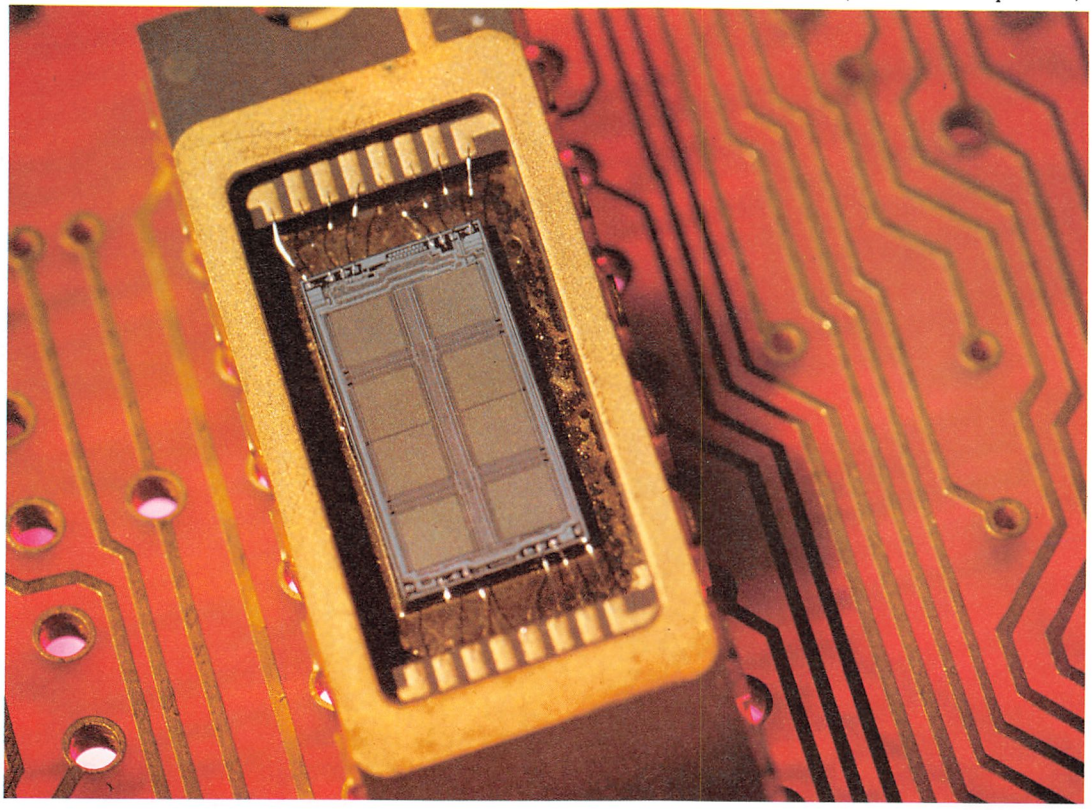
The registers that have been covered so far use inverter stages known as **ratioed inverters**. The name comes from the fact that when the input is high and the clock is high, the transistors of the inverter ($T_2$ and $T_3$ of *figure 7*) form a voltage divider between power supply voltage, $V_{DD}$, and ground. The inverter's output voltage therefore depends on the ratio of the on resistances of the transistors.

Resistances of the transistors depend largely on their physical dimensions. For example, if the channel of transistor $T_3$ is made longer and narrower than that of transistor $T_2$, then transistor $T_3$'s resistance will be greater than that of transistor $T_2$, and the logic 0 output voltage will be close to zero. There are disadvantages to this configuration: the inverter is larger than we might wish since the parasitic storage capacitance is charged through transistor $T_3$ – the higher resistance limits the register's operating speed. The ratio of a MOSFET's channel length to channel width is known as its **aspect ratio**.

To reduce the area occupied by the inverter, a different arrangement is used, in which the transistors have identical geometries and hence take up less space. *Figure 11* shows this dynamic ratioless inverter shift register. Note that no power supply voltage is used in this inverter. The clock pulses must supply the required energy to this circuit, so power dissipation is related to the clocking frequency. Let's look at how the circuit works when the input is at logic 0. During a clock pulse control voltage $V_1$ is logic 1, so transistor $T_2$ is off and transistor $T_3$ is on. Capacitor $C_3$ then charges through transistor $T_3$ to logic 1. At the end of the clock pulse, control voltage $V_1$ becomes logic 0 and so transistors $T_2$ and $T_3$ are off. The result is that capacitor $C_3$ stores the voltage across it, and an inversion has taken place.

When the input is at logic 1 and a clock pulse occurs, both transistors $T_2$ and $T_3$ are on and capacitor $C_3$ is charged to logic 1. Therefore, during the clock pulse no inversion takes place. However, after the clock pulse, control voltage $V_1$ is logic 0, transistor $T_3$ turns off, and transistor $T_2$ remains on. Now, capacitor $C_3$ discharges through transistor $T_2$ and hence inversion of the input takes place. Since this circuit absorbs power from the clock, the clock must supply the charging currents.

*(continued in part 19)*



Paul Brierley

**Left:** a 64K RAM chip.